MOLECULAR IMPRINTING • HOW PRE-AZTECS MADE MEXICO BLOOM

SCIENTIFIC AMERICAN

Immunity's Peacekeepers:
Cells That Save Us from Ourselves

OCTOBER 2006
WWW.SCIAM.COM

Catastrophysics

WHAT MAKES A STAR BLOW UP? THE MYSTERY OF A SUPERNOVA

Rolling **Ballbots**

Nanowires from **Viruses**

Global Warming and **Mass Extinctions**

Who Owns Your Body Tissues?
see page 22

october 2006

contents

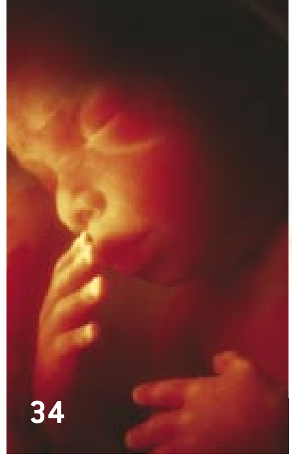
features

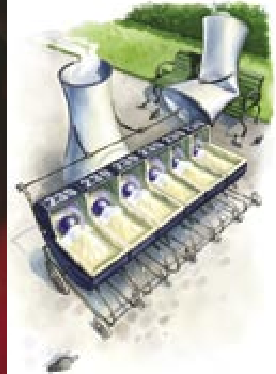
SCIENTIFIC AMERICAN Volume 295 Number 4

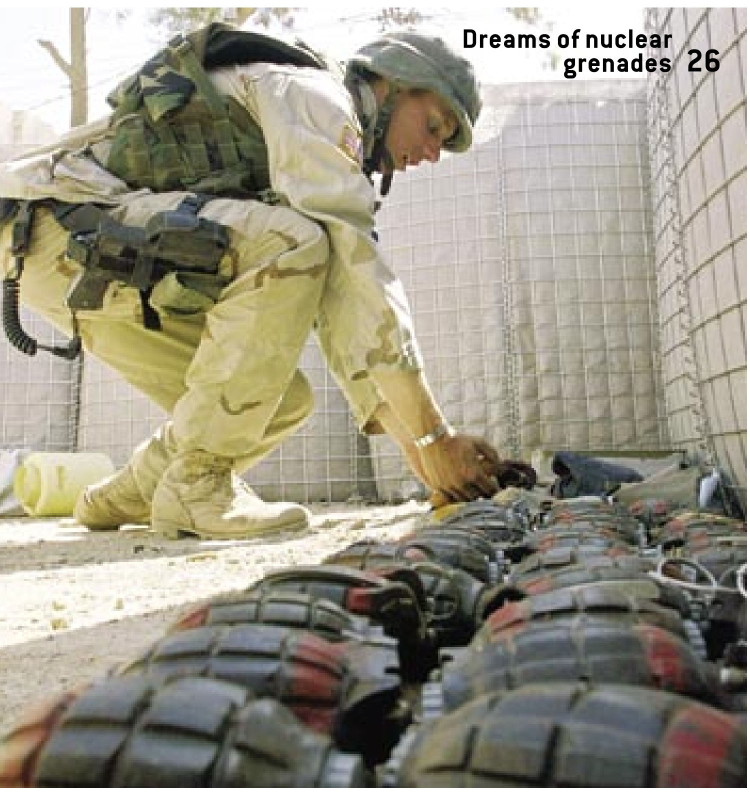
72 Like a robotic unicyclist

SCIENTIFIC AMERICAN Volume 295 Number 4

departments


34


100


Dreams of nuclear grenades 26

columns

Cover image by Phil Saunders, Space Channel Ltd.; photograph on page 3 by Brian Maranan Pineda; Tinni Chowdhury (*shadow*).

Scientific American (ISSN 0036-8733), published monthly by Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111. Copyright © 2006 by Scientific American, Inc. All rights reserved. No part of this issue may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording for public or private use, or by any information storage or retrieval system, without the prior written permission of the publisher. Periodicals postage paid at New York, N.Y., and at additional mailing offices. Canada Post International Publications Mail (Canadian Distribution) Sales Agreement No. 40012504. Canadian BN No. 127387652RT; QST No. Q1015332537. Publication Mail Agreement #40012504. Return undeliverable mail to Scientific American, P.O. Box 819, Stn Main, Markham, ON L3P 8A2. Subscription rates: one year $34.97, Canada $49 USD, International $55 USD. Postmaster: Send address changes to Scientific American, Box 3187, Harlan, Iowa 51537. **Reprints available:** write Reprint Department, Scientific American, Inc., 415 Madison Avenue, New York, N.Y. 10017-1111; (212) 451-8877; fax: (212) 355-0408. **Subscription inquiries: U.S. and Canada (800) 333-1199; other (515) 248-7684. Send e-mail to sacust@sciam.com** Printed in U.S.A.

Let There Be Light

It is practically a rite of passage that scientists who reach a certain level of eminence feel compelled to publicly announce and explain their religious beliefs. The new books by Owen Gingerich and Francis Collins, reviewed this month on page 94, follow in the footsteps of Arthur Eddington and Max Planck. Yes, these authors say, they believe in God, and no, they see no contradiction between their faith and their research—indeed, they see each as confirming the other.

SCIENCE AND FAITH can coexist happily as long as neither tries to take on the functions of the other.

Why this enduring fascination? Doubtless it is partly a reaction to the tensions that always seem to arise between science and religion: the recurring war over the teaching of evolution and creationism, the statements by physicists that they are plumbing the instant of "creation" or searching for a "God particle," the reassurances of some evangelicals that a Second Coming will make global warming irrelevant. In writing books about their own faith, religious scientists may be hoping to point the way to reconciliations for the rest of society.

Yet the tension may be greatly exaggerated. Americans are famously religious, but according to studies by the National Science Foundation, they say that they hold science in higher regard than do the people of any other industrial country. Surveys indicate that scientists are only half as likely as the general public to describe themselves as religious, but 40 percent still do. As Albert Einstein wrote, it takes fortitude to be a scientist—to persevere despite the frustrations and the long lonely hours—and religious inspiration can sometimes provide that strength.

Unquestionably, the findings of science conflict with certain religious tenets. Cosmology, geology and evolutionary biology flatly contradict the literal truths of creation myths from around the world. Yet the overthrow of religion is not a part of the scientific agenda. Scientific research deals in what is measurable and definable; it cannot begin to study what might lie beyond the physical realm or to offer a comprehensive moral philosophy.

Some scientists (and some nonscientists) are led to atheism in part by contemplation of the success of science in accounting for observable phenomena. Some find support for their spiritual beliefs in the majesty of the reality revealed by science. Others are unmoved from agnosticism. Which philosophy an individual embraces is a personal choice, not a dictate of science. Science is fundamentally agnostic, yet it does not force agnosticism even on its practitioners.

No matter how earnest their testimonies, when researchers write about their faith in God, they are not expressing a strictly scientific perspective. Rather they are struggling, as people always have, to reconcile their knowledge of a dispassionate universe with a heartfelt conviction in a more meaningful design.

As for healing a social rift, most of the debates that are commonly depicted as religion versus science are really not questions of science at all; they are disagreements among various systems of beliefs and morals. The policy fight over embryonic stem cells, for example, centers on when and how one segment of a pluralistic society should curtail the behaviors of those who hold different values. Our attention should focus not on the illusory fault line between science and religion but on a political system that too often fails to engage with the real issues.

THE EDITORS editors@sciam.com

HANS WOLF *age Fotostock*

How to Contact Us

EDITORIAL

For Letters to the Editors:
Letters to the Editors
Scientific American
415 Madison Ave.
New York, NY 10017-1111
or
editors@sciam.com
Please include your name and mailing address, and cite the article and the issue in which it appeared. Letters may be edited for length and clarity. We regret that we cannot answer all correspondence.

For general inquiries:
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8200
fax: 212-755-1976
or
editors@sciam.com

SUBSCRIPTIONS

For new subscriptions, renewals, gifts, payments, and changes of address:
U.S. and Canada
800-333-1199
Outside North America
515-248-7684
or
www.sciam.com
or
Scientific American
Box 3187
Harlan, IA 51537

REPRINTS

To order reprints of articles:
Reprint Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8877
fax: 212-355-0408
reprints@sciam.com

PERMISSIONS

For permission to copy or reuse material from *SA*:
www.sciam.com/permissions
or
212-451-8546 for procedures
or
Permissions Department
Scientific American
415 Madison Ave.
New York, NY 10017-1111
Please allow three to six weeks for processing.

ADVERTISING

www.sciam.com has electronic contact information for sales representatives of Scientific American in all regions of the U.S. and in other countries.

New York
Scientific American
415 Madison Ave.
New York, NY 10017-1111
212-451-8893
fax: 212-754-1138

Los Angeles
310-234-2699
fax: 310-234-2670

San Francisco
415-503-3630
fax: 415-437-2892

Detroit
Karen Teegarden & Associates
248-642-1773
fax: 248-642-6138

Midwest
Derr Media Group
847-615-1921
fax: 847-735-1457

Southeast and Southwest
Publicitas North America, Inc.
972-386-6186
fax: 972-233-9819

Direct Response
Special Aditions Advertising, LLC
914-461-3269
fax: 914-461-3433

Australia
IMR Pty Ltd.
Tel: +612-8850-2220
Fax: +612-8850-0454

Belgium
Publicitas Media S.A.
+32-(0)2-639-8420
fax: +32-(0)2-639-8430

Canada
Derr Media Group
847-615-1921
fax: 847-735-1457

France and Switzerland
PEM-PEMA
+33-1-46-37-2117
fax: +33-1-47-38-6329

Germany
Publicitas Germany GmbH
+49-211-862-092-0
fax: +49-211-862-092-21

Hong Kong
Hutton Media Limited
+852-2528-9135
fax: +852-2528-9281

India
Convergence Media
+91-22-2414-4808
fax: +91-22-2414-5594

Japan
Pacific Business, Inc.
+813-3661-6138
fax: +813-3661-6139

Korea
Biscom, Inc.
+822-739-7840
fax: +822-732-3662

Middle East
Peter Smith Media & Marketing
+44-140-484-1321
fax: +44-140-484-1320

The Netherlands
Insight Publicitas BV
+31-35-539-5111
fax: +31-35-531-0572

Scandinavia and Finland
M&M International Media AB
+46-8-24-5401
fax: +46-8-24-5402

U.K.
The Powers Turner Group
+44-207-592-8331
fax: +44-207-630-9922

On the Web

WWW.SCIAM.COM

UPDATED EVERY WEEKDAY

Visit **www.sciam.com/ontheweb**
to find these recent additions to the site:

NEWS

Transgenic Goat's Milk Kicks Up Immunity

Human breast milk contains valuable antibacterial enzymes that milk from dairy animals did not—until now. Researchers report that transgenic goats can successfully produce milk containing the enzyme lysozyme and that this milk exhibits an antibacterial effect when fed to young goats and pigs. The scientists hope that in the future, enhanced nonhuman milk will give an immune boost to children in the developing world, where diarrhea takes more than two million lives every year.

BLOG

Doomsday: Not as Much Fun as You'd Think

Why is doomsday so fun to think about? Maybe it fulfills the same role as the frontier once did: to allow us to envision a new world that we can shape from scratch. The world sometimes seems so confused, and individuals so powerless, that a hard reboot can start to seem downright attractive. If we're lucky, zombies, walking plants and desert car races will distract us from the mundane reality of hardscrabble farming and discovering that the meat you just ate wasn't really albatross. How would we prepare? What would we have to do to reboot civilization? It is a fascinating intellectual exercise.

PODCAST

Visit www.sciam.com/podcast to download our free weekly audio program, *Science Talk*. Recent topics include the EPA pesticide controversy, an interplanetary bicycle ride and the Coke/Pepsi espionage case.

Ask The Experts

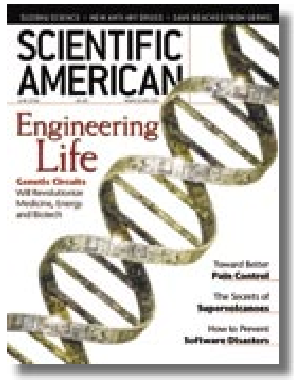
Does the moon have a tidal effect on the atmosphere as well as the oceans?

Rashid A. Akmaev, a research scientist at the University of Colorado at Boulder, explains.

CORBIS

Letters

EDITORS@SCIAM.COM



JUNE'S ISSUE held a rich vein of letter-inciting content. Software professionals challenged "Dependable Software by Design" author Daniel Jackson about the new analytical tools developed to ensure the reliability of computer programs, especially those used in critical systems. Also, the four apocalyptic volcanic eruptions in what is now California and Wyoming that each covered large areas of North America under many centimeters of ash, described in "The Secrets of Supervolcanoes," by Ilya N. Bindeman, prompted many readers to wonder how future supereruptions will affect civilization.

"Toward Better Pain Control," by Allan I. Basbaum and David Julius, documented the promise of better drug therapies to help alleviate the agony of chronic pain. Needless to say, this topic was of personal interest to many, as was evidenced by the letters requesting more information on new treatments.

DEPENDABLE DEBUGGING

In "Dependable Software by Design," Daniel Jackson clearly shows the challenges involved in developing dependable software and then loses his way describing a solution. I have been programming for 40 years. Every six months I get to sit through a presentation of a Remarkable New Product™ that will find the bugs in any piece of software (including itself). Yet they all have limitations that restrict their usefulness.

The Alloy system has at least two. First, it requires you to construct a program model in its own programming language. This just trades one programming problem for another. Second, there's no guarantee that this new model accurately reflects the real program. At the very least, the modeling constraints of Alloy need to be integrated into existing languages to be helpful.

I have no doubt that there are cases where a system like Alloy might prove its worth. But for now, the best way to get reliable software is with good programmers and hard work.

Jeff Kenton
Weston, Mass.

JACKSON REPLIES: The approach of lightweight modeling and automatic analysis isn't a panacea, but it can be a powerful tool in the software designer's arsenal.

Building and analyzing a model doesn't simply trade one programming problem for another. An Alloy model is usually orders of magnitude smaller than the program it corresponds to. It can be analyzed (using the strategy I explained) to a level of confidence currently unattainable for programs. A model can focus, unlike a program, on essential aspects of the design and is often most useful when it describes a problem alone, with no reference to putative solutions.

Because of its ability to focus on a problem and the most essential aspects of its solution, modeling can have a dramatic impact on the quality of the resulting system, even without a direct link to the code. The U.K. company Praxis, for example, has made formal modeling of requirements a centerpiece of its approach, and it routinely achieves defect rates far lower than the industry norm.

PRAY FOR SCIENCE

"No Prayer Prescription," by Christine Soares [News Scan], told of a study intended to determine the effect of third-party prayer on patient outcomes. What were they thinking? If they aren't believers, why did they bother with this study? If they are believers, did they really think the Almighty would participate in a double-blind study intended to measure the efficacy of His power?

"You shall not put the Lord your God to the test" not only reflects spiritual maturity, it makes for good science.

Geoffrey Graham
Northbrook, Ill.

REINING IN PAIN

In "Toward Better Pain Control," Allan I. Basbaum and David Julius mention the topical application of capsaicin, found in chili peppers, to relieve itching, prickling and stinging sensations. There is a "wise woman" in the village of Fortia here in Catalonia who would make up a potion of chili peppers marinated in oil. Applied regularly to aching joints, it was most effective in relieving pain. One wonders how many other such palliatives are waiting to be discovered among country folk who found them to be effective hundreds of years ago and still do in modern times.

Robin A. Flood
Catalonia, Spain

BASBAUM AND JULIUS REPLY: *It is true that many receptors involved in generating pain respond to natural products. Other readers have asked for advice on how to find a pain specialist. Many physicians will treat pain, but not all are specifically trained in pain management. Pain medicine is now a specialty recognized by the American Medical Association. To gain certification from the American Board of Pain Medicine, physicians must pass a specific exam. There are many resources for information about pain (among them: American Academy of Pain Medicine, American Pain Foundation, American Pain Society, National Pain Foundation). These organizations will not recommend a particular physician but will provide names of pain specialists in your area.*

It is always best to get a recommendation from a doctor you trust. Many pain specialists work in pain clinics, which often take a multidisciplinary approach to pain management. Note that not all pain clinics are alike; some tend to favor aggressive medical therapy, whereas others focus more on behavioral and cognitive approaches.

YUCCA AMUCK

When I saw the eruption impact map in "The Secrets of Supervolcanoes," by Ilya N. Bindeman, I was surprised that the proposed Yucca Mountain Nuclear Waste Storage facility lies close to three ancient calderas and that Yucca Mountain itself is composed in part of solidified ash flow. Congress has mandated that the facility remain secure for at least one million years. If, as you suggest, an eruption may be fairly likely over such a time period, would the ejection of this stored and extremely radioactive debris represent a potential global extinction scenario?

Ken Birman
Ithaca, N.Y.

MAGMA CUM LOUD: Mountain-size vents explode around the outer edge of an active supervolcano, producing hot gas and ash.

JULIA GREEN

BINDEMAN REPLIES: *A future supereruption in Yellowstone, Wyo., or Long Valley, Calif., indeed has a small probability of occurring. There is no reason to think, however, that such distant eruptions could damage the solidified ash beds in which nuclear waste would be stored. A supereruption near Yucca Mountain is highly improbable because the magma chamber that expelled the local ash flows crystallized long ago. Small-scale eruptions of lava and scoria, such as those that produced the 77,000-year-old Lathrop Wells basaltic cones nearby, are almost certain to occur in the next one million years, but the probability that they would intercept the Yucca site is exceedingly small.*

Furthermore, such an eruption would

Letters

become dangerously explosive only if the basaltic magma interacted with water, which is also extremely unlikely because the waste would be buried in dry ash layers that lie two kilometers above the current water table.

GRID GRIPES

In "The Science behind Sudoku," Jean-Paul Delahaye reports that "[Howard Scott] Garns died in Indianapolis in 1989 (or 1981; accounts vary) . . ." Garns was born in Indianapolis, Ind., March 2, 1905. The Social Security Death Index shows that he died in Indianapolis on October 6, 1989.

J. C. Marler
St. Louis

The chart and description on pages 84–85 show a "6" to be a forced cell in grid b. This is not correct. There is a forced 8 (not shown) in grid b in the lower right subgrid, top line.

Elizabeth From
Windham, N.H.

DELAHAYE REPLIES: Thanks to several readers for sending the date of Garns's death.

Some also wrote to report that the 6 in grid b was not forced. It is, because numbers 1, 2, 3, 4, 5, 7, 8 and 9 are in its row or in the subgrid. Moreover, the 8 is not forced at the start; it does not become forced until after forced and "only" cells have been placed.

ERRATUM In "Dependable Software by Design," by Daniel Jackson, the image on page 72 features a screen shot from avionics software by Blue Mountain Avionics. That particular software was not developed using the Alloy tool.

CLARIFICATION To answer "What are the physical and chemical changes that occur in fireworks?" [Ask the Experts], it was stated that fireworks colors are created by heating certain metals, which in turn causes electrons to jump between electron shells within atoms.

When they fall back to a lower state, they emit a photon whose wavelength determines the color. Only yellow is created this way, via the atomic emission of sodium. The other colors seen in a fireworks display are produced from burning various metallic compounds, which generate molecular emission bands.

50, 100 & 150 Years Ago

FROM SCIENTIFIC AMERICAN

Synthetic Creature ▪ Armor and Speed ▪ Power for Industry

OCTOBER 1956

WASTED RADIATION—"At present, nuclear power offers the most promising alternative to fossil fuels. However, progress in this field so far scarcely touches the heart of the problem. We speak of nuclear 'power,' but what we are really working on is nuclear heat. We are proposing to hook up the nuclear reactor to the steam turbine, an only modestly efficient invention of the 19th century, and to throw away three quarters of the energy of the nuclear reaction. It seems improvident to waste precious nuclear fuel in this fashion. Clearly the next step in power generation must be the elimination of the steam cycle and the direct conversion of radiation to electricity."

EATING MACHINE—"I call it an artificial living plant. Like a botanical plant, the machine would have the ability to extract its own raw materials from the air, water and soil [*see illustration*]. It would obtain energy from sunlight—probably by a solar battery or a steam engine. It would use this energy to refine and purify the materials and to manufacture them into parts. Then, like John Von Neumann's self-reproducing machine, it would assemble these parts to make a duplicate of itself. It could then be harvested for a material it extracted or synthesized.—Edward F. Moore"

OCTOBER 1906

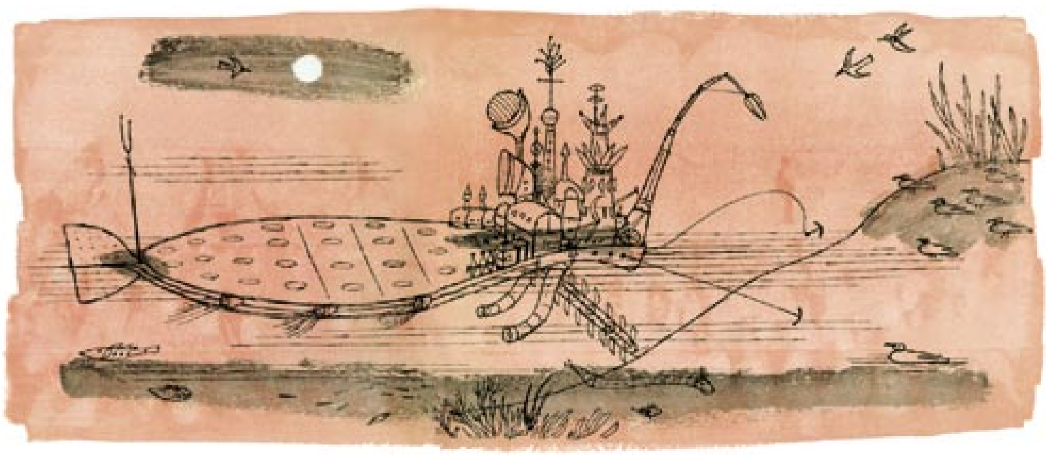
DREADNOUGHT SPEED—"The British battleship 'Dreadnought,' which has been undergoing her official trials, steamed for eight hours at an average speed of 21½ knots. We think that our naval constructors should depart from the rather conservative policy which they have followed, and allot a larger share of the displacement of our future battleships to motive power. It is true that the 'Dreadnought' is, of all foreign warships, the least likely to be arrayed against our own; but we must remember that since the mark set by this vessel will be the standard of attainment for all foreign governments, we must look for a speed of 20 knots and over in the typical battleships of the future."

BEFORE PLATE TECTONICS—"Prof. Fusakichi Omori, the eminent Japanese seismologist who has been studying the Californian earthquake, has come to the conclusion that California will be free from seismic disturbances for half a century, and probably for a much longer time. He says that the slipping of the crust of the earth was caused by the fact that at the point of weakness it was in unstable equilibrium, resulting from the redistribution of matter. It takes ages to bring this about, and the position of countless tons of matter will have to be changed, and vast quantities of earth to be carried by the rivers into the sea, before there will be so great a redistribution of matter as to cause an earthquake."

OCTOBER 1856

THIRST FOR POWER—"Most of our manufacturing towns and villages are indebted for their rise to water power. They are built on rivers and creeks where there are falls of water for driving machinery. It has now become a serious question with many manufacturers, using water power, that their supply of water is becoming more unstable every year, as the forests are cleared off. Many streams once flowing with power for the miller are now only water-worn channels. But manufactures have not decreased in our country, thanks to the power of steam. With a plentiful supply of fuel (coal), steam forms a constant trusty power for driving machinery, and a steam factory can be erected independent of rare natural localities, like water-falls. Steam factories can be conducted in or near cities and commercial marts."

SCIENTIFIC AMERICAN



SELF-REPLICATING MACHINE, in a fanciful design, gobbles minerals and plants for fuel, 1956

news SCAN

TERRORISM

A Better Defense

SHORING UP SECURITY MEANS MORE THAN HIGH TECHNOLOGY **BY DAVID BIELLO**

Visiting the White House is not unlike preparing to board a plane. Bags pass through x-ray machines and visitors through metal detectors. The White House, though, benefits from some augmentation: specifically, so-called backscatter x-ray technology, in which devices capture the radiation bounced back from objects. It can generate more detailed images than conventional x-ray machines and can even call out organic materials, such as liquid explosives.

Such technological fixes have rapidly come to the fore in the wake of recent terrorist revelations—and not for the first time: backscatter was widely touted as a solution to hijackings and other air travel incidents as long ago as the 1980s. Other high-tech solutions—millimeter-wave sensors for personal screening, quadrupole resonance to detect things like bombs in shoes, neutron-bombarding devices to precisely establish the chemical makeup of objects, as well as a host of others—can detect an array of threats, for a price. Already the Transportation Security Administration (TSA) has installed at least 100 trace detectors, designed to pick up minute quantities of various suspect compounds on clothing, luggage or people, at major airports throughout the country. But ultimately, technology can never provide the perfect security answer.

"Not a single device that anyone has built or ever will be built is 100 percent effective," explains Randy Null, chief technology officer at the TSA. "Are radiation portals in ports a good thing? The cost-benefit analysis suggests that they are," says Detlof von Winterfeldt, director of the Center for Risk and Economic Analysis of Terrorism Events at the University of Southern California. "Are surface-to-air missile deflectors? That is a lot more iffy."

This is particularly true given

TIGHT SECURITY after the August 10 foiling of a terrorist plot led to crowds and delays at international terminals, including the Leonardo Da Vinci airport in Rome.

TELENEWS/epa/CORBIS

SLIM ODDS IN THE SKY

A better understanding of the real risk of terrorism could help defeat efforts to instill fear and provoke an excessive response. Even including catastrophic events such as 9/11, the odds of the average American being a victim of terrorism in the air or elsewhere in a given year is roughly one in 500,000, experts have calculated. Comparatively, Americans face a roughly one in 6,500 chance of dying in a car accident within a given year, according to census data. "When my friends ask me, 'What can I do to travel more safely?' I tell them to drive really carefully on the way to the airport," says Brian Michael Jenkins of the Rand Corporation.

that tactics such as building bombs in midair would require tremendous levels of skill and expertise. For example, consider triacetone triperoxide (TATP), which has been employed by malefactors in the past. The necessary combination of concentrated hydrogen peroxide, sulfuric acid and acetone requires constant cooling; otherwise the liquids, when mixed, will explode weakly, enough to kill a would-be terrorist but not enough to take down a plane. And even if a bomb detonates, it might not work as planned; Aloha Airlines Flight 243—and most of its crew and passengers—survived a structurally related explosive decompression at 24,000 feet in 1988 that removed a large section of the top half of the fuselage.

A better use of limited funds includes focusing efforts on people who might choose to become terrorists. "We have not developed a strategy for the front end of the terrorist cycle: attempting to blunt the message or trying to impede recruitment," says Brian Michael Jenkins, a senior adviser at the Rand Corporation and a former member of a White House Commission on Aviation Safety and Security. "Unless we can somehow get better at these other dimensions of the conflict, then we are condemned to a strategy that is equivalent to stepping on cockroaches one at a time. We will be doing this forever."

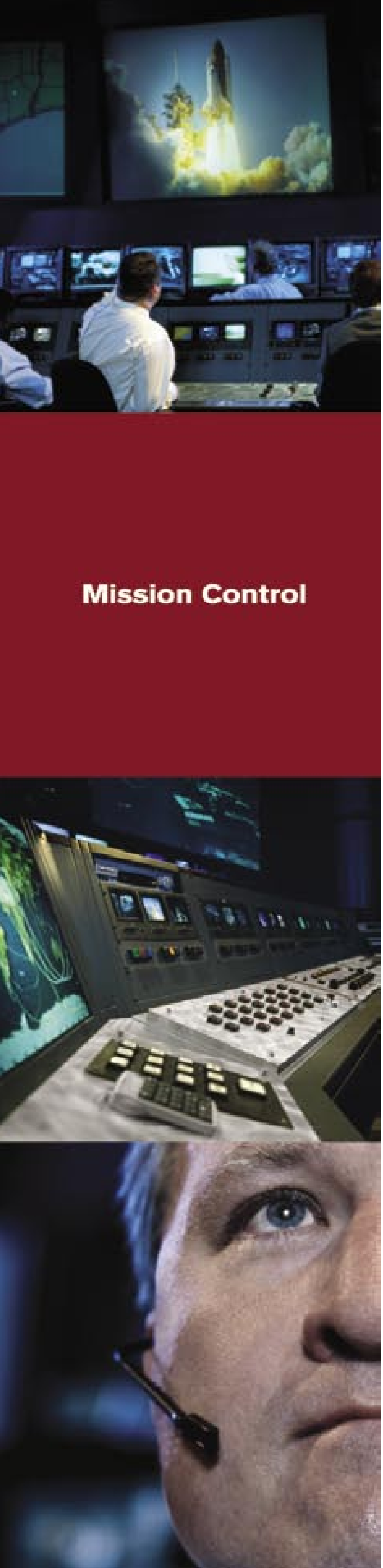
One time-tested method for addressing the front end is surveillance and infiltration, as in the case of the foiled London plot (although some experts question the timing of the arrests, arguing that further observation might have revealed additional suspects or terrorist networks). Monitoring Web sites and data mining are also critical. Once intelligence agents have collected interesting data sets, "they need to be able to connect the dots," says William Donahoo, vice president at analytical software developer Cogito, which counts the National Security Agency among its licensees. Cogito software works by linking so-called nodes—people, places or things—via their var-

ious arcs, or relationships, and then generating hypotheses based on elements such as the number of relationships a given node has. Of course, the mining is only as good as the data. "It's not meant to automatically identify bad people and have them arrested," Donahoo adds. "We're trying to help analysts do their job better."

All technology is only as good as the humans behind it: analysts, screeners, police. The TSA, for one, has pushed better training. "Over the past year we really implemented some enhanced IED [improvised explosive device] training," Null notes. "We put modular bomb kits out in the field so [officers] can get their hands on the devices and materials for bomb making." The TSA has also stepped up testing of its personnel, according to a Government Accountability Office (GAO) report, and implemented new programs that have proved effective in other countries, such as screening passengers by observation technique, or SPOT, which relies on watching and conversing with travelers who exhibit telltale fearful or deceptive behaviors.

The best defense at airports does not lie with expensive measures that focus exclusively on one tactic—bringing on board liquid explosives or shoe bombs—that can easily be changed, observes Cathleen Berrick, director of homeland security and justice at the GAO. "Build unpredictability into the screening process so the terrorists won't know what to expect," she explains. "No single layer is foolproof." Random screening with a variety of technologies could prove more effective than having all airports use the same instruments all the time. "The kinds of measures we are putting in place are now part of the permanent landscape," Rand's Jenkins notes. "We simply do not have the resources to protect everything, everywhere, all the time. That is an inherently losing battle."

A related story on explosives-detection technologies is available at www.sciam.com/ontheweb



POLICY

Uninformed Consent

DONORS REMAIN UNAWARE THEY DO NOT OWN THEIR CELLS **BY JR MINKEL**

In the summer of 2003 Washington University in St. Louis sued one of its former researchers, prostate cancer specialist William Catalona, for ownership of tissue samples collected from thousands of his patients over the years. To some of those patients, Catalona was the sole intended recipient of the specimens, and the university's actions were inconsistent with the research to which they had consented. This spring a St. Louis judge ruled in favor of the university, raising questions about the adequacy of informed consent practices.

Catalona, developer of the PSA diagnostic test for prostate cancer, had left Washington University in 2003 after conflicts over access to the samples, which the university controlled. He then sent letters to 10,000 of his patients asking that they request their samples be transferred to his new employer, Northwestern University. Six thousand patients did as he suggested. When Washington University refused their requests, some asked that their tissues be withdrawn from the repository. In response, the university anonymized the samples, removing personal information that would identify them as belonging to the donors.

After the university filed suit, eight patients joined Catalona as co-defendants. They had donated their samples specifically to him, they argued, and retained ownership of them based on their right to withdraw from research. Some of the informed consent documents they had signed did mention a right to withdraw, which derives from regulations based on the Federal Policy for the Protection of Human Subjects, known as the Common Rule. Exempted from the rule are samples that are not readily traceable to a specific person. Washington University argued that because it had anonymized the samples of those requesting withdrawal, the specimens were no longer subject to the Common Rule.

In late March, Senior District Judge Stephen N. Limbaugh in Missouri ruled the patients had gifted their samples to the university. The right to withdraw simply means that donors can stop providing samples at any time, he found. Moreover, Limbaugh wrote in his opinion, "the integrity and utility of all biorepositories would be seriously threatened if [research participants] could move their samples from institution to institution any time they wanted." The American Association of Medical Colleges expressed a similar sentiment in a brief filed in support of the university. Imagine a unique library of donated books, any of which could be removed on a whim, says David Korn, the association's senior vice president for research. "The tissue's just damned important and should be minimally encumbered," he contends.

The ruling echoes two previous major cases: *Moore v. Regents of the University of California* (1990) and *Greenberg et al. v. Miami Children's Hospital Research Institute, Inc., et al.* (2003). Patients had sued researchers for patenting a cell line or gene isolated from samples without the donors' knowledge. In both cases, courts ruled that patients had no property rights to the tissues or what was derived from them.

To patient advocates, the policy implications of the latest ruling are more important than its legal correctness. "Maintaining public trust in research is critically important. What will affect medical research is if people think they're being taken advantage of," says Ellen Wright Clayton of Vanderbilt University, who testified on behalf of Catalona's patients. "I think we need a larger discussion about control over the use of stored tissue samples for research."

The degree of control that people want over their donated samples seems to be low in most cases, but not all. In one 2005 study, investigators found that 87.1 percent of tissue donors who

filled out consent forms at the National Institutes of Health granted unlimited further use of their biological specimens, largely irrespective of the signers' age, race, residence or possibility of benefiting from research. Only 6.7 percent refused any ongoing use of their tissues, suggesting that a simple all-or-nothing consent form may be enough, the authors concluded. But if given the additional choice of being recontacted for approval of future research, a substantial minority, 26.2 percent, selected that option.

There are signs that research institutions perceive a need to maintain the public's trust. The nonprofit Coriell Institute for Medical Research in Camden, N.J., which houses the samples collected by the International HapMap Project, requires investigators to consult with the communities providing specimens if they reside in the U.S. Perhaps researchers are coming to believe that, just as with tissue, lost trust is hard to get back.

JR Minkel is a frequent contributor (of words, not cells).

SOFTWARE

Inpaint by Numbers

AN ALGORITHM TO AUTOMATE THE REPAIR OF MOVING IMAGES BY BRIE FINEGOLD

Astute moviegoers may have noticed a traveling car and its trail of exhaust in a scene from Peter Jackson's *The Lord of the Rings: The Fellowship of the Ring.* Current tools, such as Adobe Systems's After Effects "healing brush," can readily camouflage, or inpaint, small blemishes. To fix larger gaffes, though, the user must tirelessly cut small patches of image and paste them over the unwanted object. The effort typically yields mediocre results for all but the smallest repairs: on the DVD of *Fellowship,* a blurry spot is visible where the car was.

But new software using advanced mathematics may soon enable video editors to automatically inpaint a moving object quickly and seamlessly. It can even massage away large moving objects that hide other action, according to the software's principal developer, computer scientist Guillermo Sapiro of the University of Minnesota.

MOSAIC of frames identifies those that can be used to repair the video of a person walking. The damaged area is represented in the still frame (*top*) as a black rectangle.

Sapiro's knowledge of video inpainting builds on his prior experience with still photographs. In 1998 Sapiro and three colleagues observed inpainting techniques of traditional restoration artists in Paris, who started from the edges of the damage, extended the basic structure inward and then colored according to image intensity (lightness or darkness). Focusing first on grayscale images, the researchers translated the artists' techniques into mathematics (specifically, partial differential equations) that described how the intensity changed throughout the photograph in every direction.

In the case of a photograph of a bridge missing a circular patch, the program, like the artist, fills in the patch from the boundary of the circle inward. The mathematics encodes a process for simultaneously extending a special set of curves, called isophotes, along which the image intensity remains constant. The inpainting is completed when tones of gray flow along these curves from the exterior of the circle inward. "Propagating boundaries from the outside in is the common denominator among all recent inpainting techniques," Sapiro explains. For color images, the researchers calculated and then later combined the intensities of three basic colors (red, green and blue). NASA used their program to restore images of Venus (www.iua.upf.es/~mbertalmio/venus/index.html).

Extending these breakthroughs to video would seem simple—just repair each frame.

COURTESY OF GUILLERMO SAPIRO AND KEDAR A. PATWARDHAN

Despite being straightforward, this strategy might produce a choppy video or simply not work. Taking as input a frame in which the logo on a T-shirt is obscured, a photograph inpainting program would yield a perfectly fine T-shirt with no logo. But a program exploiting the temporal aspects of video could restore the logo by overlaying it with nearby frames where it was not obscured.

To incorporate time into their inpainting algorithm, the researchers viewed it as a third dimension. A circle of missing image data smoothly wandering around the film could be imagined as the result of a worm taking a wiggly path as it eats through a stack of frames. The length of the worm is the circle's duration on the film. Once these spacetime boundaries have been identified, the program inpaints the 3-D shape carved by the worm, simultaneously inpainting the 2-D circle on each frame to create a smooth result. Luckily, the user need only identify the circle in one frame, and the program will extract the entire worm by sorting static from moving parts.

Sapiro's algorithm has yet to appear in commercial software; in any case, he and his colleagues are still refining it. With Kedar A. Patwardhan of the University of Minnesota and Marcelo Bertalmio of Universitat Pompeu Fabra in Barcelona, Sapiro describes in a paper under review a faster algorithm that allows for some camera motion. Still, no software exists to remove an expanding or contracting object, such as that which occurs when a camera zooms out or in on a subject. Nor are there algorithms for completing erratic motions. As computer scientists overcome these challenges, they might even wipe away Web sites such as moviemistakes.com, where *Titanic* is listed as having 182 errors.

INPAINTING AS FLUID FLOW

The math underlying video inpainting resembles the extensively studied Navier-Stokes equations, which model the motion of fluids. Color, like liquid, flows into holes (missing circular patches on a photograph) and diffuses around obstacles (well-defined image outlines). This analogy led Guillermo Sapiro of the University of Minnesota and his colleagues to automate inpainting using techniques similar to those employed to predict fluid flow.

Examples of Sapiro's video inpainting are at www.tc.umn.edu/~patw0007/video-inpainting

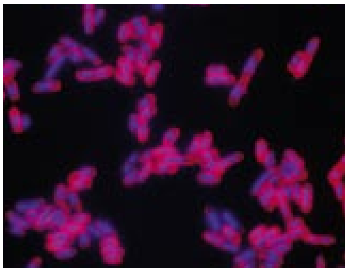
PROBIOTICS

Digestive Decoys

BACTERIA TAKE TOXIC BULLETS AIMED AT HUMAN CELLS BY CHRISTINE SOARES

Travelers to the tropics usually try to avoid consuming the local microscopic flora responsible for "Montezuma's revenge" and other, more life-threatening intestinal illnesses. But an Australian research team thinks the best way to protect against those harmful gut bacteria may be to swig more bacteria: specifically, a benign strain of *Escherichia coli* genetically engineered to absorb other bacteria's toxins.

James C. Paton and his colleagues at the University of Adelaide altered a harmless strain of *E. coli* so that it sports human-like docking sites on the surface of its cell membrane. The idea is for bacterial toxins to bind to the decoy cells instead of to cells lining the human gut. The group's latest version mimics the human cell receptors for cholera toxin, and each bug soaks up 5 percent of its own weight in poison. In a test tube, the decoy bacteria neutralize 99.95 percent of the toxin's ability to kill human cells. When a dozen baby mice were given the modified bacteria and then infected with *Vibrio cholerae,* eight survived, whereas all 12 cholera-infected control mice died. Two thirds of the test mice survived even when researchers waited four hours after infection to treat them.



TO THE RESCUE: Modified *E. coli* (*blue*) absorb Shiga toxin (*red*).

Paton has also engineered *E. coli* to bind toxins produced by more aggressive strains of its own family, including the bacterium that produces Shiga toxin and the one that often causes travelers' misery as well as lethal diarrhea among children in the developing world. He hopes his designer "probiotic" will find a niche as both an inexpensive treatment and a prophylactic in outbreak settings.

TOO CLOSE FOR COMFORT?

Concern that bacteria engineered to mimic human cells might provoke autoimmunity is based on a rare phenomenon seen after natural infections. Guillain-Barré syndrome, in which antibodies attack peripheral nerves, causes muscle weakness. A quarter of its sufferers show signs of previous infection by *Campylobacter jejuni,* which bears a surface receptor similar to one present in the myelin sheath covering human nerve fibers. Thus, antibodies developed in response to *C. jejuni* might be turning against the host's own myelin. Inflammatory factors triggered by active infection might influence antibody production, according to J. Thomas LaMont of Harvard Medical School, who thinks a passive receptor decoy therefore would probably *not* stimulate autoimmunity.

Advantages of using live bacteria against their own kind include cheap production of the beneficial bugs by fermentation and easy administration—the microbes could be swallowed along with an oral rehydration fluid, for example. Most important, because the treatment diverts toxin but does not go after the pathogen itself, it should not put pressure on the harmful bacteria to develop resistance. "We're looking for options like this, where, as a practitioner, I'd really like to have effective nonantibiotics," says J. Thomas LaMont, a gastrointestinal disease specialist at Harvard Medical School.

The concept first came to Paton and his wife, Adrienne, in the midst of a 1995 Shiga outbreak in Adelaide, where she was testing rapid diagnostic methods. "You could diagnose the children in very early stages of illness, but there was nothing you could do for them," he recalls. A week or so later they would come back with full-blown hemolytic uremic syndrome, a kidney-damaging condition caused by the toxic buildup. "So we hit upon the idea of a modified *E. coli* to mop up the toxin."

To construct his decoys, Paton inserted genes from two other bacteria types into the mild-mannered *E. coli* strain so that it would incorporate a chimeric molecule into its surface receptors. The addition of this sugar complex makes the docking sites similar enough to the human versions to bind the cholera toxin molecules. His first decoy, for Shiga toxin, turned out perfectly: all the surface molecules on the modified *E. coli* were human-receptor mimics, enabling each cell to absorb 15 percent of its own weight in toxin. The cholera toxin decoy does not generate mimic receptors quite as prolifically, but that is not the biggest hitch in developing it for use in people. Because its chimeric receptor is so structurally similar to the human type, it could, in theory, trigger an autoimmune reaction.

That is something to investigate further, Paton acknowledges, but he is more immediately worried that his bugs will run afoul of international restrictions on transgenic organisms. A solution to both potential problems may be to kill the beneficial bacteria, Paton explains: "It works when it's dead. Not quite as well, but once it's dead, it's no longer a 'genetically modified organism.'"

DEFENSE

Far-Out Physics

BIG BUDGETS KEEP "FRINGE" PROJECTS ALIVE BY DANIEL G. DUPONT

Antimatter weapons. Psychic teleportation. "Voice of God" microwaves that beam speech into the heads of unsuspecting enemies. Hand grenades that pack a nuclear punch.

To most scientists, these ideas are too far-out to warrant near-term focus. Despite their objections, these and other kinds of "fringe" science projects have taken root over the years in U.S. government agencies, delivering false hopes of coming superweapons to policymakers unschooled in the laws of physics.

Take the so-called hafnium bomb, also known as the isomer bomb. According to its proponents, this futuristic

NUCLEAR DUDS? According to a questionable theory about "isomer" bombs, grenades might someday pack kiloton yields.

AHMAD AL-RUBAYE *AFP/Getty Images*

weapon would harness the immense energy in subatomic particles known as isomers to deliver kiloton yields in a tiny package. Others have said isomers could lead to advances in powerful lasers.

Isomer-based weapons concepts have been around for decades. The basic notion is that a nuclear isomer (an element that has some excited protons) could be made to decay and release that energy, which could trigger the fusion of other atoms. The idea did not get serious attention until a leading proponent claimed in 1998 to have employed a dental x-ray machine to "trigger" the release of energy from isomers of the element hafnium.

Researchers roundly decried the results as questionable, faulty or even impossible. They could not be replicated even with much more powerful lasers, for one thing; critics also said that even if triggered, hafnium would not make for much of a weapon. At best, the process could produce only a radiological, or "dirty," bomb. But that did not stop the military from giving the idea serious consideration—and money. The Pentagon has spent at least $10 million for the hafnium bomb alone.

Sharon Weinberger, a veteran defense reporter, traces the development of such weaponry in her book, *Imaginary Weapons: A Journey through the Pentagon's Scientific Underworld,* released in June. She describes the military's pursuit of a hafnium bomb—which continues today—as "a story about the depths of self-deception and the willingness of government officials to believe in threats that don't exist," as well as fantastic weapons to counter them.

One reason why the Pentagon pursues fringe topics, according to Weinberger and others, comes down to the sheer size of the military budget: at almost $500 billion a year, it funds untold numbers of research efforts, and it is easy for lesser ideas to slip in under the radar. Another reason: congressional earmarking, in which lawmakers insert pockets of money into appropriations bills to do constituents a favor. The funds attract little oversight but

can keep alive even the most questionable projects.

Moreover, many legislators and staff "really don't understand these technologies," comments Philip Coyle, a former top Pentagon weapons inspector who now works for the Center for Defense Information, a Washington, D.C., watchdog group. That ignorance, he adds, can lead to "pretty far-out projects."

Mix in a healthy dose of secrecy, and the conditions for fringe science to flourish greatly improve, says Steven Aftergood, who runs the Project on Government Secrecy for the Federation of American Scientists. In the post-9/11 security environment, more research is being classified as secret, allowing a very few people to give the green light to and manage such projects. "Secrecy insulates funders from independent scrutiny and from the embarrassment that would follow disclosure of some of these programs," he says.

As an example of secrecy's effect, Aftergood points to a 2004 U.S. Air Force study of teleportation, or what he calls "*Star Trek*–style transportation." The air force took a lot of criticism for funding what is "pretty much universally accepted as physically untenable," Aftergood says, adding that the embarrassment came only after the project was disclosed. "Had the process been more transparent," he argues, "the taxpayers might have been spared the expense."

The air force defended that expense by claiming the need to turn over every rock, just in case. This kind of justification does not sit well with critics such as physicist Steven Koonin of the California Institute of Technology. He served on a panel that reviewed the hafnium bomb for the Pentagon in the late 1990s and says in Weinberger's book: "It's not enough to say it's 'out of the box,' it's got to be plausibly out of the box."

Daniel G. Dupont, a longtime contributor, edits InsideDefense.com and InsideGreenBusiness.com

PUBLIC HEALTH

news SCAN

Contentious Calculation

CONTROVERSY OVER CHERNOBYL'S FUTURE CANCER TOLL **BY JOHN DUDLEY MILLER**

The Chernobyl nuclear plant in Ukraine exploded 20 years ago, but the disaster will continue for another 60 years in the form of slow deaths from cancers. The accident released a plume that dropped radioactive particles throughout the Northern Hemisphere. No one has pinned down the expected toll—estimates range from thousands to tens of thousands, revealing disagreements in the way the figures should be calculated and limitations in current knowledge about radiation damage.

The most commonly reported figure is 4,000 deaths, which derives from a 2005 United Nations press release. Curiously, it called the 4,000 a "new" number from a study by "an international team of more than 100 scientists"—even though the cited work was from 1996 and was authored by only seven scientists. "Certainly the 1996 paper was never meant to make the headlines of the newspapers 10 years later," remarks lead author Elisabeth Cardis of the World Health Organization's International Agency for Research on Cancer in Lyon, France.

Stranger still, the 1996 study had estimated 9,000 deaths, not 4,000. "It was either a deliberate omission or a simply outrageous error," says David Marples, a historian at the University of Alberta who has written several books about Chernobyl. "In either case, it is not scientifically acceptable." Keith Baverstock, the former head of the European office of the WHO's radiation protection division, calls it "scientifically unacceptable" as well as "selective and misleading."

A staffer at the International Atomic Energy Agency (IAEA) strongly disagreed, arguing that both the IAEA and the WHO believe that 4,000 is correct, because the calcu-

GENETIC CONSEQUENCES

Cancer deaths may not be Chernobyl's only legacy. Timothy Mousseau, a biologist at the University of South Carolina, cites 25 different studies that show increased genetic mutation in plants, animals and children in contaminated areas. Researchers have also observed more spontaneous abortions in people who were highly contaminated as well as increases in birth defects, developmental disabilities and heart ailments in their children. Although no conclusive evidence exists, Mousseau thinks that the possibility that the mutations caused these problems is "very likely."

QUESTION OF MULTITUDES: A young cancer patient undergoes treatment in Kiev in April. The number of cancer deaths that may result from Chernobyl is hotly contested.

ODED BALILTY *AP Photo*

lation for the additional 5,000 is so uncertain that both organizations "considered it insignificant." But a WHO spokesperson says that to correct "the imprecisions" of the 2005 U.N. release, the organization issued a statement in April announcing that "there may be up to 9,000 excess deaths due to Chernobyl." The IAEA's charter "to accelerate" the expansion of nuclear power worldwide biases it "toward underreporting Chernobyl deaths," says Robert Alvarez, a former Department of Energy senior policy adviser.

Still, 9,000 deaths may be a vast underestimate. Cardis had confined her 1996 analysis to contaminated areas of Ukraine, Belarus and western Russia. Sixty-four percent of Chernobyl's radiation fell outside the former Soviet Union, according to a 1988 U.N. report.

This past April, Cardis and a new team extended the original study to all of Europe. On top of 2,200 cleanup-worker deaths the 1996 study estimated, they predict that through 2065 another 6,700 to 38,000 European residents (with a "best guess" of 16,000) might succumb to Chernobyl-caused cancers. Other calculations, including one by Baverstock and another by the European Green Party, peg the worldwide death toll at 30,000 to 60,000.

Chernobyl death estimates are also controversial because they are based on data from survivors of the 1945 atomic bombings of Japan. Those individuals received large doses all at once; Chernobyl exposed people to small, continuous doses for years. Uncertainty about the exact amounts of radiation that the Japanese victims received has contributed to the broad ranges in Chernobyl estimates. Also, some investigators believe that any amount of radiation, no matter how small, harms the body and that the damage is proportional to the dosage. Others think that a threshold exists below which radiation is harmless.

A threshold effect suggests that the estimates of tens of thousands of deaths, from calculations that assume that harm is linearly proportional to dosage, are well off the mark. "They're not people," remarks Antone Brooks, a radiation biologist at Washington State University. "They're numbers generated from a hypothesis." Brooks believes a threshold exists, so he thinks that the IAEA was correct to drop the 5,000 predicted deaths in less contaminated areas.

Who is right in the Chernobyl guesstimate game may never be known. Cancer kills one out of every four people, meaning that about 117 million Europeans will die of non-Chernobyl cancers through 2065, Cardis remarks. So even though the nuclear accident may ultimately cause many thousands of deaths, Chernobyl cancers will, unfortunately, be impossible to detect.

John Dudley Miller, a writer based in Cleveland, was a U.S. Navy submarine nuclear engineering officer.

ASTRONOMY

Fat Side of the Moon

The moon's far side bulges at its equator, a peculiarity that has long puzzled investigators. Scientists conjectured that the bulge formed when magma oceans that covered the young moon solidified while deformed by gravity and lunar spin, but the hypothesis failed to match theories of the moon's early orbit with its precise dimensions. Researchers at the Massachusetts Institute of Technology now calculate that they could explain the bulge if the moon's orbit 100 million to 200 million years after the moon formed was about twice as close to Earth and more oval. This orbit, resembling present-day Mercury's in completing three rotations for every two revolutions, would have proved ideal to help freeze the bulge in place. The findings, appearing in the August 4 *Science,* also suggest a time on Earth when the moon cycled through its phases in just 18 hours and raised tides four times a day at up to 10 times the strength.

—*Charles Q. Choi*

PHYSICS

Bubble Adhesion

When two smooth, slick surfaces are submerged in water and brought within 100 nanometers of each other, something odd happens: they adhere, even though they are too far apart for electrostatic forces to bridge them. C. Jeffrey Brinker of Sandia National Laboratories and his colleagues repeated this experiment with rougher silica surfaces that repel water more than smooth ones do. Using a special kind of atomic force microscope to control the spacing of the surfaces, they found that this strange attraction kicked in at an even longer distance of up to two microns and that it was accompanied by the formation of a vapor bubble in between. The surfaces' antagonism toward liquid water creates a partial vacuum that draws them together, the group concludes in the August 3 *Nature.*

—*JR Minkel*

CANCER

Fido's Transmissible Tumor

Dogs have a sexually transmitted cancer apparently spread via tumor cells that roam the world as parasites. Robin Weiss of University College London and his colleagues investigated canine transmissible venereal tumor, a malignancy found in the domestic dog and possibly in relatives such as the gray wolf and the coyote. Their DNA analysis of dog tissue samples from five continents reveals that the tumor cells did not belong to the dogs in which they festered. Instead they were all nearly genetically identical to one another and similar to cells from wolves or closely related dog breeds from China or Siberia. Judging by the number of mutations the cells accumulated, the cells' origin dates to sometime between 200 and 2,500 years ago, making them the oldest known cancer lineage. This cancer could yield clues on how tumors survive and elude the immune system, the investigators write in the August 11 *Cell*. —*Charles Q. Choi*

CANINE CONDITION of venereal tumors is the oldest known mammalian cancer lineage.

TAXI/GETTY IMAGES

BRIEF POINTS

- **Mast cells, the source of histamine in allergy attacks, release an enzyme that breaks down viper and honeybee venom, rather than simply exacerbating the effects of those toxins as previously assumed.**

  Science, July 28

- **Internet search engines could in theory produce unwanted feedback, making popular sites even more popular. Fortunately, however, the specificity and diversity of searches keep the popularity bias in check.**

  Proceedings of the National Academy of Sciences USA, August 22

news SCAN

DATA POINTS: DUST DEVIL

The debris and combustion particles resulting from the destruction of the World Trade Center on September 11, 2001, have compromised lung function in surviving rescue workers. Medical researchers are studying members of the Fire Department of New York City (FDNY), who are tested regularly to see how much air they can exhale through a tube.

Number of exposed FDNY rescue workers: 11,766

Number who arrived before or during the collapse: 1,660

Two days later: 8,185

Three or more days later: 1,921

Median liters of air workers could exhale in one second during their last test before 9/11: 4.30

During their first test after 9/11: 3.93

Percent who had below-normal capacity before 9/11: 6.8

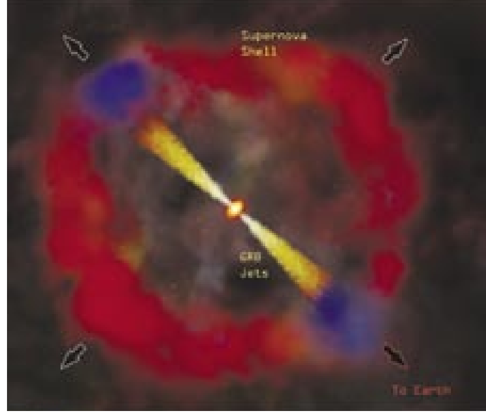
After 9/11: 15.3

Equivalent years of aging for lungs, one year after 9/11: 12

SOURCE: American Journal of Respiratory and Critical Care Medicine, *August 2006*

COSMOLOGY

That Way Lies Confusion



COSMIC MYSTERY involves galaxy counts between quasars and gamma-ray bursts, such as GRB 020813, in which gas from a supernova surrounds a black hole producing two particle jets (*artist's rendition*).

Astronomers know that if they point their telescopes at quasars, they will spy an average of one "foreground" galaxy in front of every fourth quasar. Because the universe is uniform, the number of foreground galaxies should be the same for, say, a group of observed gamma-ray bursts. Only it is not. In a paper that is generating considerable buzz among the stellarati, Jason X. Prochaska of the University of California, Santa Cruz, finds an average of about one foreground galaxy for each of 15 bursts. If the result holds, then astronomers are misinterpreting a key aspect of foreground gas—potentially posing a serious cosmological problem, because they use the gas to estimate the composition of the earliest galaxies and the distribution of dark matter, which makes up 90 percent of the matter in the universe. Foreground galaxies might be unexpectedly dusty, obscuring some quasars; they might be focusing light from gamma-ray bursts, causing astronomers to miss fainter ones. Or the supposed galaxies might be gas from the bursts themselves, the researchers note in the September 20 *Astrophysical Journal Letters*. *—JR Minkel*

TRANSGENIC CROPS

Cotton-Picking Results

Cotton genetically modified to produce a pesticide from the microbe *Bacillus thuringiensis* does not guarantee long-term financial benefit. Cornell University researchers and their colleagues interviewed 481 Chinese farmers in five major cotton-growing provinces. For years Bt farmers there reaped at least a third more profit than conventional growers from the money saved by slashing pesticide use up to 70 percent. But then Bt-resistant insects, such as mirids, proliferated (though for at least one season, an unusually cool, wet summer led to a large mirid outbreak). The new pests forced the farmers to spray crops as much as conventional growers did. That ate away the profits, because Bt seed costs three times more than conventional seed. The scientists, who presented their findings at the July 25 American Agricultural Economics Association meeting in Long Beach, Calif., suggested targeting secondary pests with natural predators or modifying cotton even further to resist them.

—Charles Q. Choi

HEALTH

Womb Woes

The so-called fetal origins hypothesis predicts that poor health in utero should be followed by chronic disease in adulthood, and indeed studies of brief famines back up the claim. But an analysis by Columbia University economist Douglas Almond indicates that problems for the less robust unborn extend to socioeconomic success, too. He zeroed in on people who were prenatally exposed to influenza during the 1918 pandemic. Detailed census data from the 1960s to the 1980s show that members of this group were up to 15 percent less likely to graduate from high school, had among men 5 to 9 percent lower wages (because of disability), and were 15 percent more likely to be poor compared with siblings and others of their generation, Almond reports in the August *Journal of Political Economy*. Policies aimed at improving prenatal health could accordingly have a strong effect on future earnings, he notes. *—JR Minkel*

FETAL HEALTH sets the tone for adult well-being.

M. WEISS/NASA/CXC (*top*); NEIL BROMHALL *Photo Researchers, Inc.* (*bottom*); ILLUSTRATION BY MATT COLLINS

news
SCAN

BY THE NUMBERS

To Serve Man

THE SERVICES SECTOR PROVES TO BE A KEY TO PROSPERITY BY RODGER DOYLE

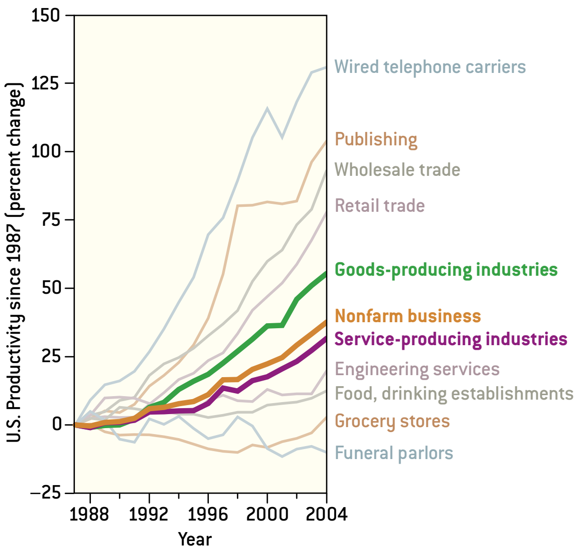
Weak productivity growth became a major concern in the U.S. beginning in the late 1970s, when overall productivity in the country slowed considerably. Researchers attributed most of the slowdown to the service sector, where more than eight out of 10 Americans now work. Services, economists believed, were not amenable to rapid productivity growth.

The work of two economists decisively puts this notion to rest. Using data from the Bureau of Economic Analysis and the Bureau of Labor Statistics, Jack E. Triplett and Barry P. Bosworth of the Brookings Institution constructed measures of overall productivity for the goods- and service-producing industries. Their data, indicated by the thick lines on the chart, show that productivity in the service sector has risen considerably since 1995 and at a much faster rate than from 1970 to 1995.

Industries that made the largest investments in computer hardware and software in the 1980s and early 1990s showed the largest productivity gains after 1995. The long-distance telephone industry, for example, invested heavily in information technology after AT&T broke up in 1984. Telephone companies subsequently saw strong productivity growth. The use of scanners, computers and the Internet has contributed to the strong productivity growth in retail trade, although some businesses, such as grocery stores, have failed to advance in recent years. Railroads, trucking and warehousing have had declining productivity, apparently reflecting their low investment in infotech. Some traditional industries, such as funeral homes, may have had no need to incorporate the latest communications technology. Industries that modernized have helped to slow inflation in the years since 1995.

The U.S. is well ahead of Europe in productivity gains, in part because of the lack of impediments. Zoning laws in France, for example, protect food shops from competition; in Germany, small state-owned and cooperative retail banks have little incentive to consolidate and modernize because they are not answerable to the capital markets. Economic growth has been strongest in those countries that have deregulated the most.

Productivity growth is the key to income growth, but according to economists Ian Dew-Becker and Robert Gordon of Northwestern University, the benefits of the recent surge in productivity went mainly to those in the top 10 percent of the income distribution, with the top 1 percent gaining disproportionately. The failure of workers to benefit from the productivity surge is not entirely clear, but it is probably related to a federal minimum wage that has not kept up with inflation.



Productivity in the economic sectors shown in the chart can be measured because they have an input (hours worked) and an output (sales). This is not true of noncommercial activities, including public education and government-backed health care, where the output cannot be measured in sales. Comparative measures of productivity in these sectors, which would have to be calculated in a different way, would be most useful as a guide to public policy.

Rodger Doyle can be reached at rodgerpdoyle@verizon.net

RODGER DOYLE; SOURCES: BARRY P. BOSWORTH *Brookings Institution (goods- and service-producing industries, nonfarm business)*; BUREAU OF LABOR STATISTICS *(all other data)*

FAST FACTS: U.S. WORKFORCE

In 2005, 141,218,000 Americans worked full-time. The private sector employed 83 percent, which breaks down as:

GOODS-PRODUCING: 17.4
Manufacturing: 10.1
Construction: 5.4
Agriculture: 0.6
All other: 1.3

SERVICE-PRODUCING: 65.6
Retail trade: 11.2
Health care, social assistance: 10.6
Accommodations, food: 7.8
Administrative, waste management: 5.8
Professional, technical: 5.3
Finance and insurance: 4.3
Wholesale trade: 4.1
Transportation and warehousing: 3.1
Information: 2.2
Education: 2.1
Real estate: 1.6
All other: 7.5

FURTHER READING

Productivity in the U.S. Services Sector. Jack E. Triplett and Barry P. Bosworth. Brookings Institution, 2004.

Where Did the Productivity Growth Go? Ian Dew-Becker and Robert J. Gordon. National Bureau of Economic Research Working Paper No. 11842, December 2005.

Is the 21st Century Productivity Expansion Still in Services? And What Should Be Done about It? Barry P. Bosworth and Jack E. Triplett. National Bureau of Economic Research, July 2006.

Darwin on the Right

Why Christians and conservatives should accept evolution

By MICHAEL SHERMER

According to a 2005 Pew Research Center poll, 70 percent of evangelical Christians believe that living beings have always existed in their present form, compared with 32 percent of Protestants and 31 percent of Catholics. Politically, 60 percent of Republicans are creationists, whereas only 11 percent accept evolution, compared with 29 percent of Democrats who are creationists and 44 percent who accept evolution. A 2005 Harris Poll found that 63 percent of liberals but only 37 percent of conservatives believe that humans and apes have a common ancestry. What these figures confirm for us is that there are religious and political reasons for rejecting evolution. Can one be a conservative Christian and a Darwinian? Yes. Here's how.

Calling God a watchmaker is belittling.

1. Evolution fits well with good theology. Christians believe in an omniscient and omnipotent God. What difference does it make *when* God created the universe—10,000 years ago or 10,000,000,000 years ago? The glory of the creation commands reverence regardless of how many zeroes in the date. And what difference does it make *how* God created life—spoken word or natural forces? The grandeur of life's complexity elicits awe regardless of what creative processes were employed. Christians (indeed, all faiths) should embrace modern science for what it has done to reveal the magnificence of the divine in a depth and detail unmatched by ancient texts.

2. Creationism is bad theology. The watchmaker God of intelligent-design creationism is delimited to being a garage tinkerer piecing together life out of available parts. This God is just a genetic engineer slightly more advanced than we are. An omniscient and omnipotent God must be above such humanlike constraints. As Protestant theologian Langdon Gilkey wrote, "The Christian idea, far from merely representing a primitive anthropomorphic projection of human art upon the cosmos, systematically repudiates all direct analogy from human art." Calling God a watchmaker is belittling.

3. Evolution explains original sin and the Christian model of human nature. As a social primate, we evolved within-group amity and between-group enmity. By nature, then, we are cooperative and competitive, altruistic and selfish, greedy and generous, peaceful and bellicose; in short, good and evil. Moral codes and a society based on the rule of law are necessary to accentuate the positive and attenuate the negative sides of our evolved nature.

4. Evolution explains family values. The following characteristics are the foundation of families and societies and are shared by humans and other social mammals: attachment and bonding, cooperation and reciprocity, sympathy and empathy, conflict resolution, community concern and reputation anxiety, and response to group social norms. As a social primate species, we evolved morality to enhance the survival of both family and community. Subsequently, religions designed moral codes based on our evolved moral natures.

5. Evolution accounts for specific Christian moral precepts. Much of Christian morality has to do with human relationships, most notably truth telling and marital fidelity, because the violation of these principles causes a severe breakdown in trust, which is the foundation of family and community. Evolution describes how we developed into pair-bonded primates and how adultery violates trust. Likewise, truth telling is vital for trust in our society, so lying is a sin.

6. Evolution explains conservative free-market economics. Charles Darwin's "natural selection" is precisely parallel to Adam Smith's "invisible hand." Darwin showed how complex design and ecological balance were unintended consequences of competition among individual organisms. Smith showed how national wealth and social harmony were unintended consequences of competition among individual people. Nature's economy mirrors society's economy. Both are designed from the bottom up, not the top down.

Because the theory of evolution provides a scientific foundation for the core values shared by most Christians and conservatives, it should be embraced. The senseless conflict between science and religion must end now, or else, as the Book of Proverbs (11:29) warned: "He that troubleth his own house shall inherit the wind." SA

Michael Shermer is publisher of Skeptic *(www.skeptic.com). His new book is* Why Darwin Matters.

BRAD HINES

BRUCE GILBERT/EARTH INSTITUTE

Fiddling while the Planet Burns

Will the *Wall Street Journal*'s editorial writers accept a challenge to learn the truth about the science of global climate change? By JEFFREY D. SACHS

Another summer of record-breaking temperatures brought power failures, heat waves, droughts and tropical storms throughout the U.S., Europe and Asia. Only one place seemed to remain cool: the air-conditioned offices of the editorial board of the *Wall Street Journal*. As New York City wilted, the editors sat insouciant and comfortable, hurling editorials of stunning misdirection at their readers, continuing their irresponsible drumbeat that global warming is junk science.

Now, I have nothing against the *Wall Street Journal*. It is an excellent paper, whose science column and news reporting have accurately and carefully carried the story of global climate change. The editorial page sits in its own redoubt, separated from the reporters—and from the truth.

A July 14 editorial, "Hockey Stick Hokum," epitomizes the problem. The climate-change "hockey stick" is a graph first published in 1998 by Michael Mann et al. that attempted to reconstruct the mean surface temperature on the planet during the period A.D. 900 to the present. The conclusion of the study was that we are now in that interval's warmest range of temperatures, therefore adding support to the overwhelming evidence that man-made climate change is already well under way.

The newspaper gets a free pass on editorial irresponsibility.

The *Wall Street Journal* editorial page for years has railed against these scientific findings on climate change, even as the global scientific consensus has reached nearly 100 percent, including the reports commissioned by the skeptical Bush White House. Thus, the hockey stick became the bête noire of the editorial page as well as of the dwindling "climate skeptic" community, and right-wing officials such as Representative Joe Barton of Texas, chairman of the House Subcommittee on Energy and Air Quality, took up the attack.

In response to these growing political pressures, the National Research Council of the National Academy of Sciences conducted a major independent scientific review and updating of the hockey stick data and analysis. It came down squarely on the side of the Mann study, stating that "the committee finds it plausible that the Northern Hemisphere was warmer during the last few decades of the 20th century than during any comparable period over the preceding millennium." Significant uncertainties remain for global temperatures before 1600, the NRC noted, but it emphasized that "surface temperature reconstructions for periods prior to the industrial era are only one of multiple lines of evidence supporting the conclusion that climatic warming is occurring in response to human activities, and they are not the primary evidence."

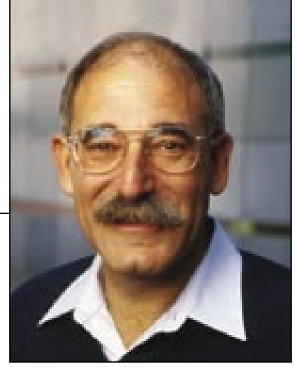
The *Wall Street Journal* editorial page completely ignored this report. Instead it cited one commissioned by Barton from three statisticians with no background in climate science, who quibbled with aspects of Mann's methodology. Nevertheless, on this flimsy and misleading basis, the editorial page declared that "there's no reason to believe that Mr. Mann, or his 'hockey stick' graph of global temperature changes, is right," called the research "dubious" and said that the climate science community "often more closely resembles a mutual-admiration society than a competitive and open-minded search for scientific knowledge."

The *Wall Street Journal* is the most widely read business paper in the world. Its influence is extensive. Yet it gets a free pass on editorial irresponsibility. The Earth Institute at Columbia University has repeatedly invited the editorial team to meet with leading climate scientists. On many occasions, the news editors have eagerly accepted, but the editorial writers have remained safe in their splendid isolation.

Let me make the invitation once again. Many of the world's leading climate scientists are prepared to meet with the editorial board of the *Wall Street Journal* and to include in that meeting any climate skeptics that its editorial board wants to invite. The board owes it to the rest of us to conduct its own "open-minded search for scientific knowledge." SA

An expanded version of this essay is available online at www.sciam.com/ontheweb

Jeffrey D. Sachs is director of the Earth Institute at Columbia University.



WEBB CHAPPELL

The New Age of Wireless

Technologies that turn broadcasting "bugs" into features that open radio spectrum to novel uses will be a boon for consumers By ANDREW LIPPMAN

Before 1968 no one in the U.S. could connect anything to the AT&T telephone system unless Western Electric, AT&T's manufacturing arm, provided it. The Federal Communication Commission's landmark "Carterfone" decision erased that policy and ignited an explosion of communications innovations, including faxes, fast modems, PBXs, burglar alarms, answering machines and phone mobility. Although AT&T no longer owned the whole pie, the slice that it kept became part of a far larger industry.

That same explosive growth is beginning in wireless mobile. Microprocessors are now so fast that they can synthesize and handle directly both sound-and-image data and radio signals. Meanwhile the emergence of agile, end-to-end networks is creating unprecedented opportunities in what for 100 years have been staid communications structures. No matter what you think of the wireless devices you have today, you ain't seen nothing yet. Radio is just getting interesting.

Spectrum need not be a finite resource.

Mobile phones will become programs loaded into whatever physical "engine" is convenient or perhaps into many at once. Instead of a "family" phone service plan, you might someday have a "molecular" account that makes accessible any radio-accessible thing or data that you choose. You could decide whether to put your dog's view of the world online, whether to monitor your blood sugar level from afar or whether to talk to someone through your eyeglasses. Broadband will become the province of a person rather than a wire to a home.

The broadcast radio spectrum, which has typically been regarded as limited and interference-plagued, will become open and accessible everywhere by anything. Of course, some broadcasters and registered networks will still rely on keeping certain airwaves empty and silent, and they will be used by legacy devices that we are loath to discard, such as cell phones and AM radios. But grander possibilities await radios that cooperatively sense one another's proximity, use one another to economize on radiated energy and battery life, and turn ever more remote regions of the spectrum into fertile territory for personal use.

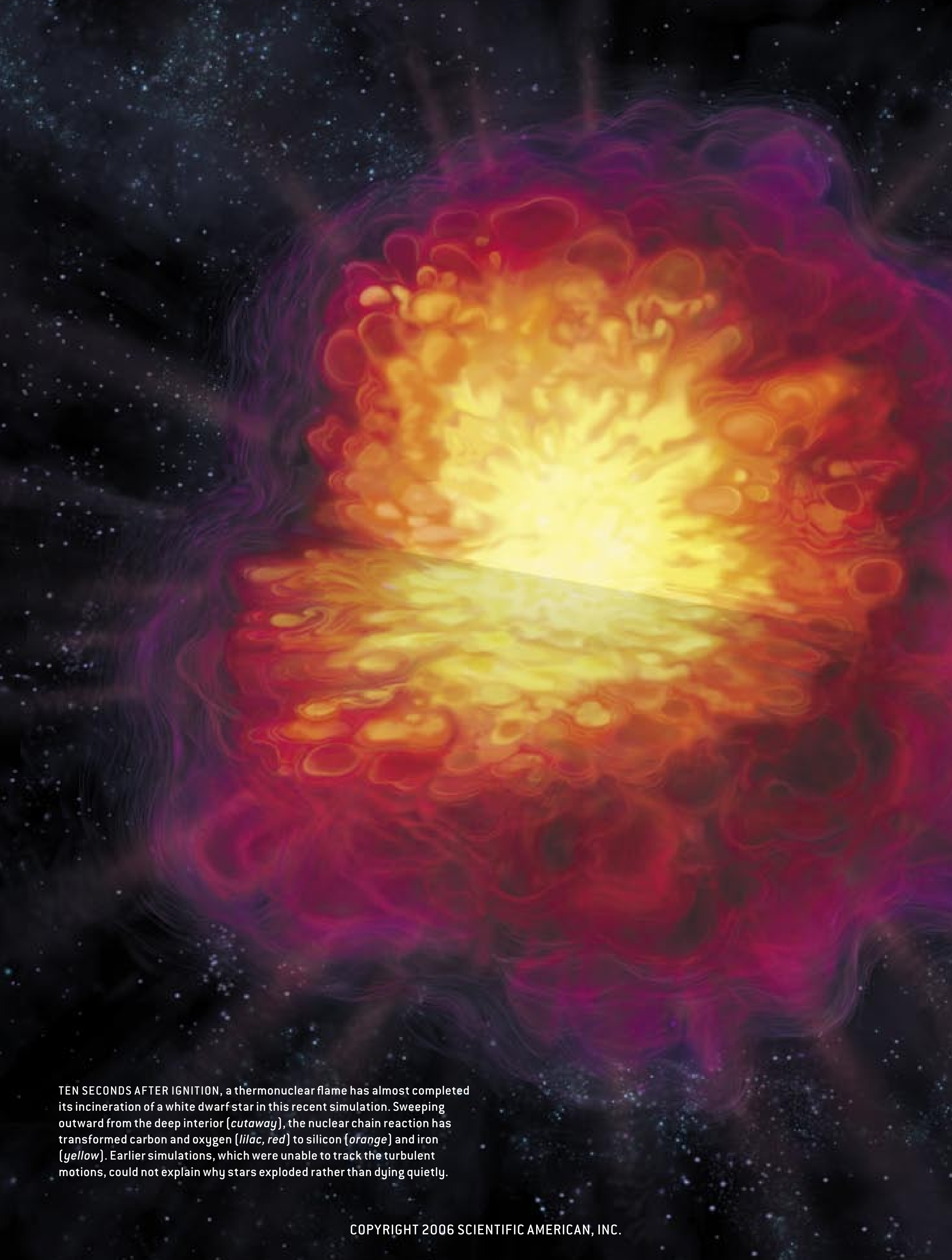
Disparate demonstrations paint this new picture of wireless communications. For example, the multipath phenomenon, in which buildings and walls bounce multiple copies of a signal to a receiver, was once just the source of ghosts in television pictures. But in essence, those reflectors are also sending additional energy that would have been lost. Thus, they can also be regarded as independent transmitters. Multiple-input, multiple-output radios built to take advantage of that effect can improve communications.

Other work has built ad hoc networks of mobile radios that at each moment dynamically select for intermediate relays requiring the least energetic connections. One radio might momentarily be in a dead spot, but another will be in a hot spot for passing on a communication. As radios become cheaper than their batteries, adding transmitters becomes more efficient than adding power to make reliable systems. Even more important, the dichotomy between broadcasting and point-to-point connectivity disappears; the two work together by design.

A new discipline called network coding uses broadcasting to save bandwidth by coding and then relaying bits for more than one receiver. Decoding involves combining several transmissions, including your own. Using this principle, we have built a demonstration telephone system in which the default is that everyone can hear anyone—a wireless party line. We call it "push to listen" because you decide how loudly and in which ear you would like to place any speaker's voice. Stock traders, emergency workers and perhaps conference callers might find it particularly useful.

The broadcast nature of wireless is thus a feature rather than a bug: it can save energy, increase efficiency and nurture new ideas. And spectrum need not be regarded as a fixed and finite resource to be divided among users. Instead it can support more communication as more communicators use it. The theory has been in place for a few years, but now it is becoming real. The pie will soon start to grow, and there will be enough slices for all. SA

Andrew Lippman heads the Media Lab's Viral Communications program and co-directs the Communications Futures program at the Massachusetts Institute of Technology.



TEN SECONDS AFTER IGNITION, a thermonuclear flame has almost completed its incineration of a white dwarf star in this recent simulation. Sweeping outward from the deep interior (*cutaway*), the nuclear chain reaction has transformed carbon and oxygen (*lilac, red*) to silicon (*orange*) and iron (*yellow*). Earlier simulations, which were unable to track the turbulent motions, could not explain why stars exploded rather than dying quietly.

How to Blow Up a Star

By Wolfgang Hillebrandt,
Hans-Thomas Janka
and Ewald Müller

It is not as easy as you would think. Models of supernovae have failed to reproduce these explosions—until recently

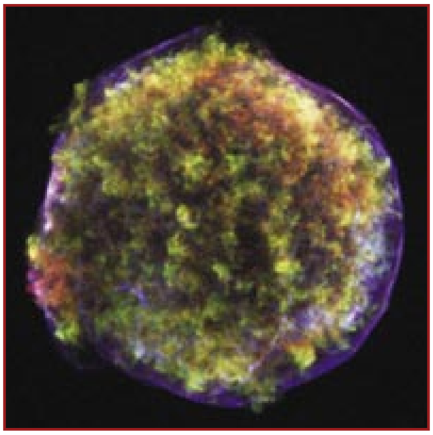
On November 11, 1572, Danish astronomer and nobleman Tycho Brahe saw a new star in the constellation Cassiopeia, blazing as bright as Jupiter. In many ways, it was the birth of modern astronomy—a shining disproof of the belief that the heavens were fixed and unchanging. Such "new stars" have not ceased to surprise. Some 400 years later astronomers realized that they briefly outshine billions of ordinary stars and must therefore be spectacular explosions. In 1934 Fritz Zwicky of the California Institute of Technology coined the name "supernovae" for them. Quite apart from being among the most dramatic events known to science, supernovae play a special role in the universe and in the work of astronomers: seeding space with heavy elements, regulating galaxy formation and evolution, even serving as markers of cosmic expansion.

Zwicky and his colleague Walter Baade speculated that the explosive energy comes from gravity. Their idea was that

a normal star implodes until its core reaches the density of an atomic nucleus. Like a crystal vase falling onto a concrete floor, the collapsing material releases enough gravitational potential energy to blow the rest of the star apart. An alternative emerged in 1960, when Fred Hoyle of the University of Cambridge and Willy Fowler of Caltech conceived of the explosions as giant nuclear bombs. When a sunlike star exhausts its hydrogen fuel and then its helium, it turns to its carbon and oxygen. Not only can the fusion of these elements release a titanic pulse of energy, it produces radioactive nickel 56, whose gradual decay would account for the months-long afterglow of the initial explosion.

Both these ideas have proved to be right. Of the supernovae that show no signs of hydrogen in their spectra (designated type I), most (type Ia) appear to be thermonuclear explosions, and the rest (types Ib and Ic) result from the collapse of stars that had shed their outer hydrogen layers. Supernovae whose spectra include hydrogen (type II) are thought to arise from collapse as well. Both mechanisms reduce an entire star to a shell of gaseous debris, and gravitational collapse events also leave behind a hyperdense neutron star or, in extreme cases, a black hole. Observations, notably of supernova 1987A (a type II event), have substantiated this basic theoretical picture [see "The Great Supernova of 1987," by Stan Woosley and Tom Weaver; Scientific American, August 1989].

Even so, explaining supernovae is still a major challenge for astrophysicists. Computer simulations have had trouble reproducing the explosions, let alone their detailed properties. It is reassuringly hard to get stars to explode. They regulate themselves, remaining very stable for millions or billions of years. Even dead or dying stars have mechanisms causing them to peter out rather than blowing up. Figuring out how these mechanisms are overcome has taken multidimensional simulations that push computers to, and beyond, their limits. Only very recently has the situation improved.



TYCHO'S SUPERNOVA, a thermonuclear explosion observed by renowned Danish astronomer Tycho Brahe in 1572, left behind a cloud of silicon, iron and other heavy elements glowing in x-rays (*green, red*). The shock front (*thin blue shell*) expands outward at 7,500 kilometers a second.

Overview/*Supernovae*

- By all rights, stars should be stable, sober creatures, preferring to die with a whimper than with a bang. Astronomers have struggled to understand why some of them go supernova. These explosions have resisted efforts to describe them in a simplified way, making them perhaps the most complex phenomena in all of astrophysics.
- Theorists have gradually ratcheted up the sophistication of their models and have recently succeeded at last in reproducing the two main types of supernovae. The key has been to capture all three spatial dimensions in fine enough detail to track the turbulent flow dynamics.
- They have discovered that the explosions can be seriously lopsided, stirring the debris (which includes newly synthesized chemical elements) thoroughly. In the kind of explosion that leaves behind a neutron star, this star recoils and careens across the galaxy at high speed.

Blowing Up Is Hard to Do

IRONICALLY, the stars that are thought to blow up as type Ia supernovae are usually paragons of stability—namely, white dwarf stars. A white dwarf is the inert remnant of what used to be a sunlike star. If left unmolested, it stays more or less in the state it was born, gradually cooling down and fading out. But Hoyle and Fowler argued that if a white dwarf tightly orbits another star, it may accrete matter from its companion, grow in mass and become ever more compressed at its center, until it reaches densities and temperatures sufficiently high to explosively fuse carbon and oxygen.

The thermonuclear reactions should behave rather like an ordinary fire. A front of burning should sweep through the star, leaving behind a pile of nuclear ash (mainly nickel). At any moment, the fusion reactions would occur in a tiny volume, most likely at the surface of ash-filled bubbles floating in the deep interior of the white dwarf. Because of their lower density, the bubbles would be buoyant and try to rise toward the surface of the star—much like steam bubbles in a pot of boiling water.

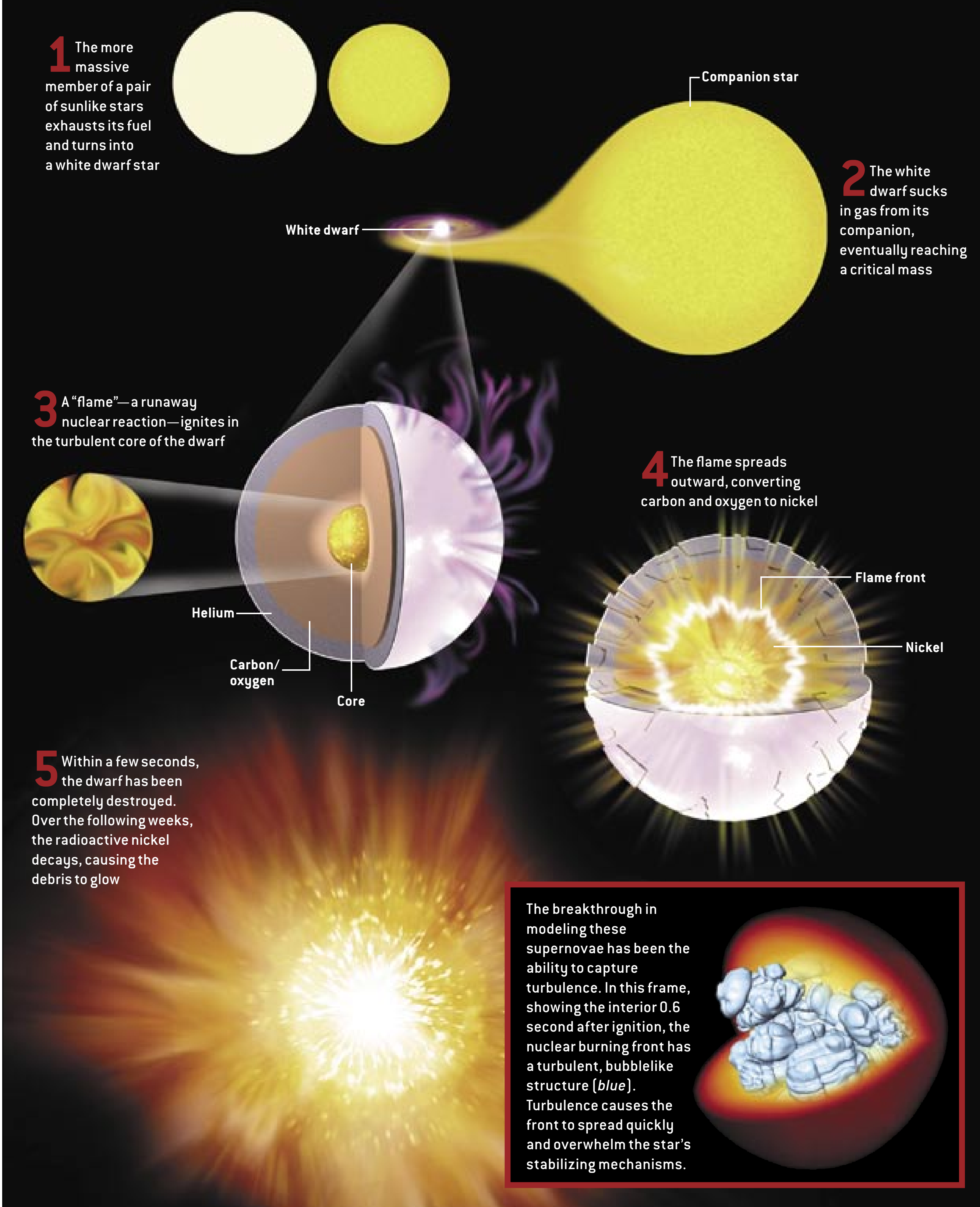
The trouble with this picture was that the thermonuclear flame should fizzle; the energy released would cause the star to expand and cool, thereby quenching the burning. Unlike an ordinary bomb, a star has no walls to confine it and prevent this self-extinguishment.

Coupled with this theoretical stumbling block has been a practical one. No lab experiments that reproduce the conditions in supernovae can be performed, and observations are subject to their own limitations. The best approach that astrophysicists have is to try to simulate the explosion on a computer. That is a momentous task. Currently the most precise simulations, which our group has conducted using an IBM p690 supercomputer, divide the star into a grid of up to 1,024 elements on a side, capturing details as small as a few kilometers across. A single run requires a few times 10^{20} arithmetic operations, and for such a complex problem, the supercomputer can perform several 10^{11} operations a second. All in all, it takes al-

PAT RAWLINGS SAIC (*preceding pages*); NASA/CXC/JESSICA WARREN, JOHN P. HUGHES ET AL. *Rutgers University*

THERMONUCLEAR SUPERNOVA

One class of supernova, known as type Ia, is the sudden nuclear detonation of an entire star.



DON DIXON (*illustrations*); WOLFGANG HILLEBRANDT AND FRIEDRICH ROEPKE *Max Planck Institute for Astrophysics* (*simulation*)

most 60 processor-years. The computational tricks that simplify simulations in other areas of science do not apply to supernovae, which involve highly asymmetrical flow, extreme conditions, and a huge range of spatial and temporal scales. Particle physics, nuclear physics, fluid dynamics and general relativity are complicated enough on their own, but a supernova simulation must consider all of them at once.

Under the Hood

THE SOLUTION has come from an unexpected quarter: the physics of car engines. Mixing and igniting gasoline and oxygen in an engine generates turbulence. Turbulence, in turn, increases the surface area of the flames by wrinkling and stretching them. The rate of fuel consumption, proportional to flame area, goes up. A star, too, is naturally turbulent. Because the gas moves vast distances at high velocities, even a slight disturbance quickly turns a smooth flow into a turbulent one. In a supernova, the rising hot bubbles should stir up the material, causing the nuclear burning to spread so fast that the star has no time to compensate.

In a properly working internal-combustion engine, the flame propagates at a subsonic speed limited by the rate at which heat diffuses through the material—a process called deflagration. In an engine suffering from knocking, the flame propagates at a supersonic rate, driven by a shock wave passing through the fuel-oxidizer mixture and compressing it—a process known as detonation. Thermonuclear flames can spread in the same two ways. Detonation, being the more violent, would incinerate the star more thoroughly, leaving behind only the most highly fused elements, such as nickel and iron. Observationally, however, astronomers detect a wide variety of elements in these explosions, including silicon, sulfur and calcium. That pattern suggests the nuclear burning propagates, at least initially, as deflagration.

CRAB NEBULA is the gaseous debris of a core-collapse supernova observed in 1054. At the center is a neutron star (*arrow*) spewing particles that cause the gas to glow (*blue*). The outer filaments consist mostly of hydrogen and helium from the disrupted massive star.

NASA/ESA/ALLISON LOLL AND JEFF HESTER *Arizona State University*

In the past few years, thermonuclear deflagrations have finally been convincingly modeled by our group as well as teams at the University of California, Santa Cruz, and the University of Chicago. The computer code we have honed draws on methods developed for the study of chemical combustion and even of weather. Turbulence is inherently a three-dimensional process. In a turbulent cascade, kinetic energy shifts from large length scales to small ones, where ultimately it dissipates as heat. In other words, the flow becomes ever more finely patterned. Therefore, the simulations have to be three-dimensional, too. That has become feasible only very recently.

Simulating supernovae in their full dimensionality has revealed complex mushroomlike structures—hot bubbles rising in a layered fluid, wrinkled and stretched by turbulence. The increase of the fusion reaction rate wrought by turbulence leads to the disruption of the white dwarf in just a few seconds. The debris expands at about 10,000 kilometers a second, in fair agreement with what the observations show.

That said, many open questions linger. What initially ignites the white dwarf is not understood at all. Another problem is that deflagration should eject a large fraction of the white dwarf essentially unchanged, whereas observations suggest that at most a small part of the star remains unaltered. So the explosion cannot be a pure deflagration; some detonation must be involved. Theorists have yet to explain why both processes would operate. Nor can they account for the observed variety of explosions. It may well be that accretion onto a white dwarf is not the only way to set off a type Ia supernova; a merger of two white dwarfs might, too.

Gravity's Tomb

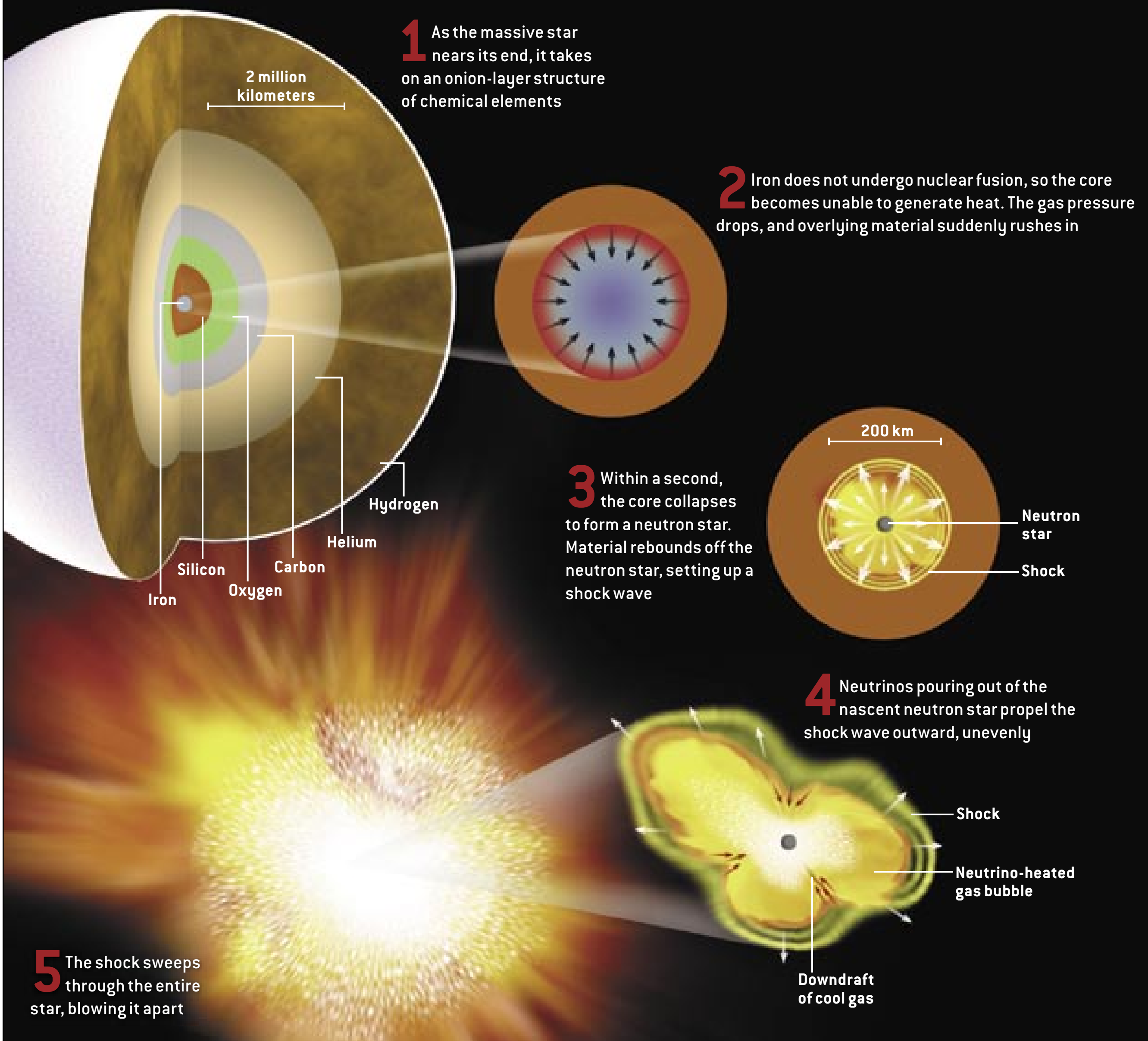
THE OTHER MAJOR TYPE of supernova, arising from the collapse of a stellar core, is even trickier to explain. Observationally, these events come in a wider variety than thermonuclear supernovae do: some have hydrogen, some do not; some explode into a dense interstellar environment, some into nearly empty space; some eject large amounts of radioactive nickel, some do not. The range of explosive energy and expansion velocity is enormous. The most powerful produce not only the classic supernovae blasts but also long-duration gamma-ray

THE AUTHORS

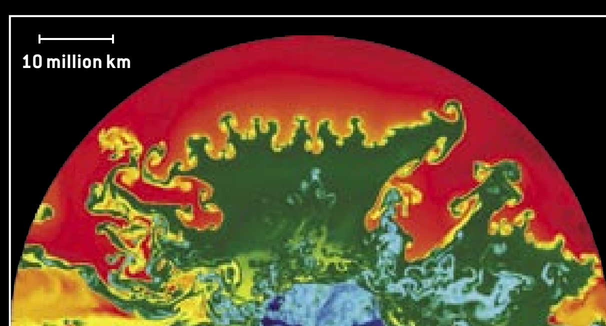
WOLFGANG HILLEBRANDT, HANS-THOMAS JANKA and *EWALD MÜLLER* are scientists at the Max Planck Institute for Astrophysics (MPA) in Garching, Germany, and teach at Munich Technical University. Hillebrandt is one of the three directors of the MPA. His main research areas are nuclear astrophysics, stellar evolution and supernovae explosions, which he says he got into because he was fascinated by "really big explosions." He won the Physics Prize of the German Physical Society in 1982 for his work on rapid neutron-capture nucleosynthesis. In winter he can often be found on the ski slopes and in summer on a sailboat. Janka is interested in neutrino astrophysics, neutron star evolution, and supernovae and gamma-ray bursts. One month after he had started his Ph.D. project, Supernova 1987A was detected, and his career (let alone the universe) was never the same again. He spends his spare time painting, drawing and doing handicrafts. Müller is leader of a research group on numerical and relativistic astrophysics. Together with Janka, he won the Heinz Billing Award for Scientific Computing in 1993. His fascination with astrophysics was inspired by science-fiction novels. He is still a big fan of sci-fi movies and enjoys hiking in the Bavarian Alps.

CORE-COLLAPSE SUPERNOVA

The other class of supernova involves the implosion of a star at least eight times as massive as the sun. This class is designated type Ib, Ic or II, depending on its observed characteristics.



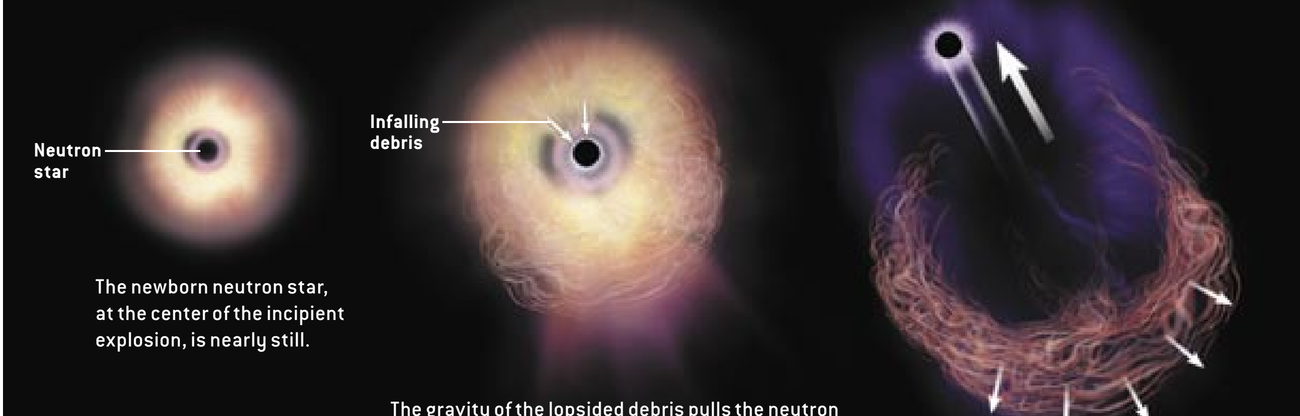
Recent simulations have made huge progress in tracking the chaotic motions during the explosion. In this frame, showing the interior five and a half hours into the explosion, large rising bubbles have helped drive the shock wave a distance of 300 million kilometers. Neutrinos, usually an antisocial breed of particle, stream out of the initial implosion in such quantities and with such high energies that they play a decisive role. The turbulence mixes carbon, oxygen, silicon and iron from deep down (*blue, turquoise*) into the overlying helium (*green*) and hydrogen (*red*).



DON DIXON (*illustrations*); KONSTANTINOS KIFONIDIS *Max Planck Institute for Astrophysics* (*simulation*)

THE SUPERNOVA ROCKET EFFECT

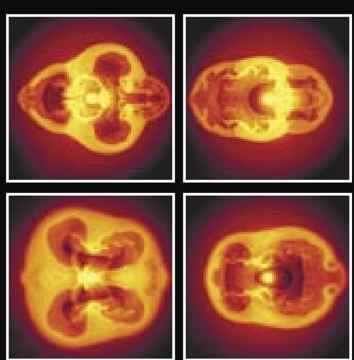
Observers have puzzled over why neutron stars zip through the galaxy at high speed. The new models of core-collapse supernovae offer an explanation based on the intrinsic asymmetry of these explosions.



The newborn neutron star, at the center of the incipient explosion, is nearly still.

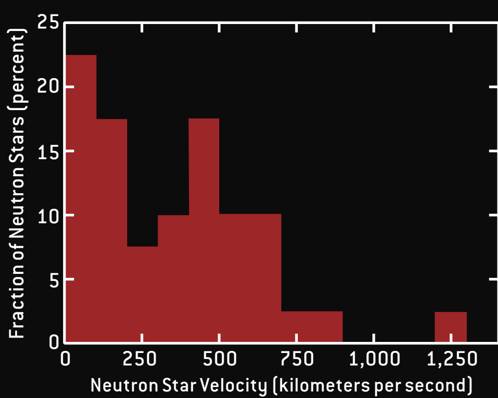
The gravity of the lopsided debris pulls the neutron star in a certain direction, and debris falling onto the neutron star provides an additional kick.

These forces eject the neutron star. (Because of the overall momentum balance, the neutron star is pulled in the direction from which the infalling debris is coming.)



According to simulations, the asymmetry develops early in the explosion. Tiny differences in the way stars die lead to large variations in the degree of asymmetry.

These variations, in turn, lead to a range of neutron star velocities. Comparing the predicted velocities to observations can provide a test of the models.



DON DIXON (*top illustrations*); LEONHARD SCHECK *Max Planck Institute for Astrophysics* (*simulations*); TINNI CHOWDHURY (*graph*)

bursts [see "The Brightest Explosions in the Universe," by Neil Gehrels, Luigi Piro and Peter J. T. Leonard; Scientific American, December 2002]. This heterogeneity is only one of many longstanding puzzles. Core-collapse supernovae are the prime candidates for the origin of the heaviest of all elements, such as gold, lead, thorium and uranium, which can be created only under special conditions. But nobody knows whether such conditions are indeed realized when stellar cores implode.

Although the basic idea of core collapse sounds simple—the collapse releases gravitational binding energy that blasts out material—the details are hard to grasp. By the end of its life, a star with more than about 10 solar masses has developed an onionlike structure, comprising layers with successively heavier elements. The core is composed mainly of iron, and its structural integrity is maintained by quantum repulsion between electrons. Eventually, though, the weight of the star overwhelms the electrons. They get squeezed into the atomic nuclei, where they react with the protons to form neutrons and electron neutrinos. The neutrons and remaining protons, in turn, get packed closer and closer until their own repulsive forces come into play, stopping the collapse.

At this point, the implosion somehow reverses and becomes a powerful outflow. Matter diving deep into the gravitational well is lifted out again. In the classic theory, this task is achieved by the shock wave that is set up as the outer stellar layers crash with supersonic speed onto the abruptly decelerated inner core. This shock wave moves outward, compressing and heating the material it encounters.

The trouble is that the shock uses up its energy and eventually stalls. Simulations show that the energy of the implosion quickly dissipates. So how does the star blow itself apart?

The germ of an answer emerged in pioneering work by Stirling Colgate and Richard White in 1966 and in more modern computer simulations by Jim Wilson in the early 1980s. (All three worked at what is now known as Lawrence Livermore National Laboratory.) They suggested that the shock wave is

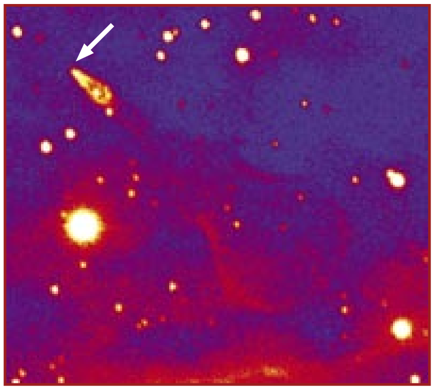
not the only way that energy from the core can reach the outer layers of the star. Maybe the neutrinos generated in the collapse play a role. At first, the idea sounds strange: neutrinos are notoriously unsociable; they typically interact with other particles so weakly that they are difficult even to detect. But in a collapsing star, they are endowed with more than enough energy to drive an explosion—and in the extremely dense conditions, they couple to matter more strongly. They heat a layer around the inner core of a supernova, raising the pressure behind the stalled shock wave.

Rocket Science

IS THAT EXTRA PUSH enough to revive the shock, drive it outward and complete the explosion? Computer simulations of the process indicated that it was not. Although the gas absorbs neutrinos, it also emits them, and the models suggested that the losses would dominate and the explosions would fizzle out. These models, however, made a radical simplification: they assumed the star to be spherically symmetrical. Thus, they ignored critical multidimensional phenomena, such as convection and rotation, which are clearly important because observed supernovae leave behind highly aspherical, jumbled debris.

This realization seems to be the key to solving the supernova problem. Multidimensional simulations show that the plasma in the neutrino-heated layer around the inner core of a supernova develops buoyant bubbles and mushroom-like plumes. Convection carries energy to the shock wave, pushing it farther out and triggering an explosion.

The new picture has very appealing implications. When the explosion sets in relatively slowly, the bubbles of hot, expanding plasma, separated by downflows of cooler matter, have time to merge together. Eventually the flow pattern consists of just a few or even a single rising bubble surrounded by downdrafts. The consequence is a lopsided explosion, explaining why supernova remnants are so skewed. Another asymmetry is that the stalled shock front can deform, causing the explosion to develop an hourglass shape. Additional flow instabilities occur when the revived shock rushes outward and passes through the layers of the progenitor's onion-shell structure. The chemical elements synthesized during the star's life and in the explosion event get mixed together.



GUITAR NEBULA is a shock wave set off by a neutron star (*at arrow*) zipping through gas at 1,600 kilometers a second. The explosion that created the star must have been seriously lopsided to fling it to such a speed.

S. CHATTERJEE AND J. CORDES *Cornell University*

Because the stellar debris is ejected with more power to one side, the neutron star at the middle is kicked in the opposite direction, just as a skateboard skids away when you jump off it. Our group has found recoil velocities of more than 1,000 kilometers a second, which matches the observed motions of most neutron stars. Some neutron stars move more slowly, which suggests that the bubbles in the explosion that created them did not have time to merge. A unified picture thus emerges, in which a variety of phenomena stem from just one underlying effect.

Despite considerable progress over the past years, however, no existing model has yet reached sufficient realism to demonstrate how these supernovae work in their full glory. All the models still involve approximations and simplifications. A full model would have seven dimensions: space (in three coordinates), time, neutrino energy, and neutrino velocity (described by two angular coordinates). Moreover, it would allow for all three types, or flavors, of neutrino. Around the world, a major effort is on to develop new computer hardware and software to achieve such a model.

One of researchers' many aims is to study whether explosions might be triggered in more than one way. Magnetic fields, for example, might tap the rotational energy of the newly formed neutron star, giving the shock wave an extra push. Magnetic fields might also squeeze matter outward along the rotational axis in two polar jets. Such effects might explain the most powerful explosions. Gamma-ray bursts, in particular, appear to involve jets moving almost at the speed of light. The core of such a burst may collapse not to a neutron star but to a black hole.

As modelers make progress, observers, too, are poking into barely explored realms, looking not just for electromagnetic radiation but also for neutrinos and gravitational waves. The collapsing stellar core, its violent boiling at the onset of explosion, and its possible transition to a black hole not only produce an intense burst of neutrinos but also shake the fabric of spacetime. Unlike light, which is heavily processed by the overlying layers, these signals escape directly from the cataclysmic abyss at the center of the explosion. New neutrino and gravitational-wave detectors may be the source of our next surprise in the saga of how stars die. SA

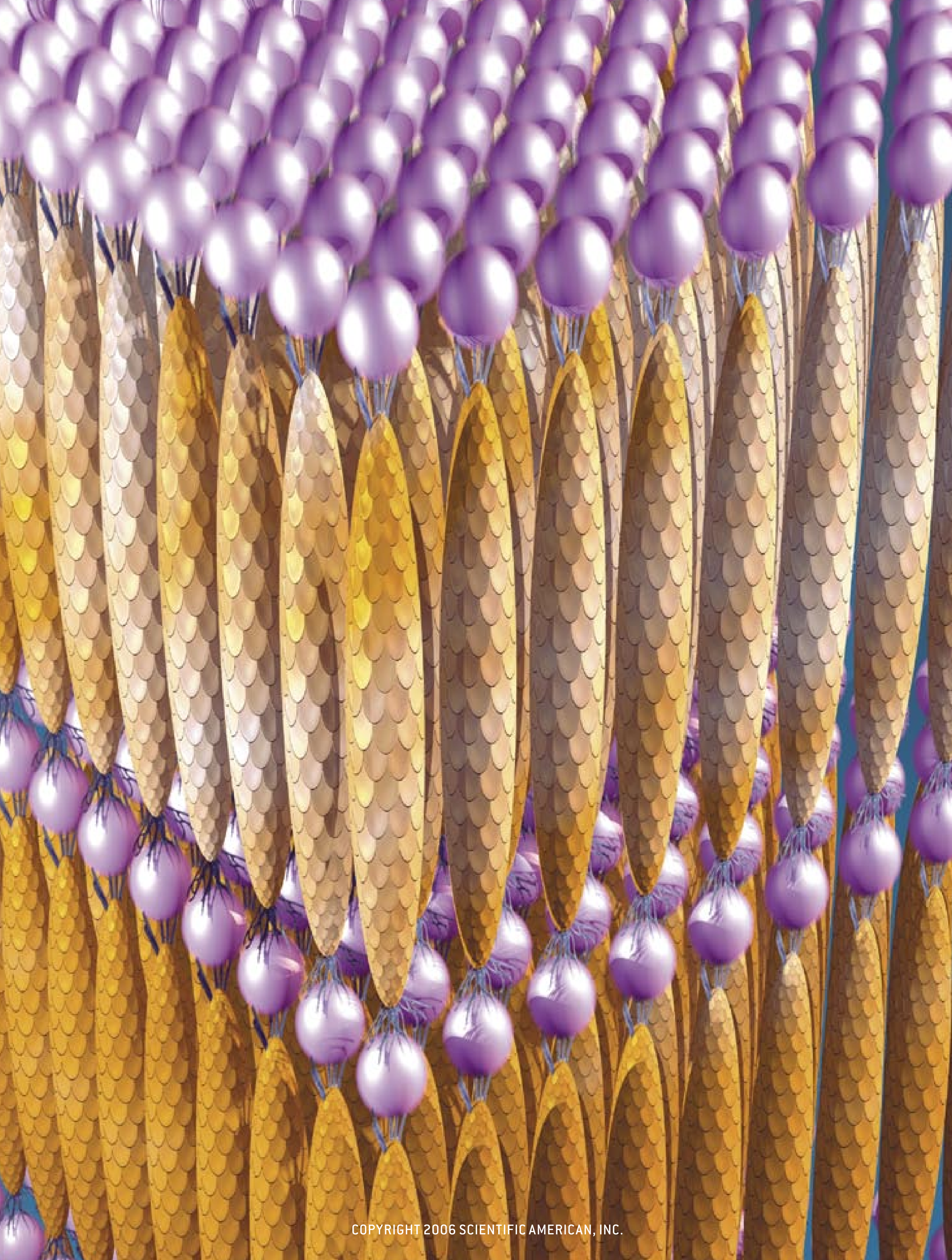
MORE TO EXPLORE

Supernova Explosions in the Universe. A. Burrows in *Nature,* Vol. 403, pages 727–733; February 17, 2000.

Full-Star Type Ia Supernova Explosion Models. F. K. Röpke and W. Hillebrandt in *Astronomy and Astrophysics,* Vol. 431, No. 2, pages 635–645; February 2005. Preprint available at **arxiv.org/abs/astro-ph/0409286**

The Physics of Core-Collapse Supernovae. S. Woosley and H.-Th. Janka in *Nature Physics,* Vol. 1, No. 3, pages 147–154; December 2005. Preprint available at **arxiv.org/abs/astro-ph/0601261**

Multidimensional Supernova Simulations with Approximative Neutrino Transport. L. Scheck, K. Kifonidis, H.-Th. Janka and E. Müller in *Astronomy and Astrophysics* (in press). Preprint available at **arxiv.org/abs/astro-ph/0601302**



VIRAL NANO ELECTRONICS

M.I.T. breeds viruses that coat themselves in selected substances, then self-assemble into such devices as liquid crystals, nanowires and electrodes

By Philip E. Ross

For many years, materials scientists wanted to know how the abalone, a marine snail, constructed its magnificently strong shell from unpromising minerals, so that they could make similar materials themselves. Angela M. Belcher asked a different question: Why not get the abalone to make things for us?

She put a thin glass slip between the abalone and its shell, then removed it. "We got a flat pearl," she says, "which we could use to study shell formation on an hour-by-hour basis, without having to sacrifice the animal." It turns out the abalone manufactures proteins that induce calcium carbonate molecules to adopt two distinct yet seamlessly melded crystalline forms—one strong, the other fast-growing. The work earned her a Ph.D. from the University of California, Santa Barbara, in 1997 and paved her way to consultancies with the pearl industry, a professorship at the Massachusetts Institute of Technology, and a founding role in a start-up company called Cambrios in Mountain View, Calif.

Belcher had a grand plan to develop biological agents that could move molecules around like so many bricks, building structures from the ground up, a strategy known in the nanotechnology world as self-assembly. But to begin the task, she needed a critter more tractable than the abalone, a high-maintenance animal that grows slowly and is something of a one-note Charlie. She wanted something small, spry and flexible—a cross between Maxwell's famous molecule-sorting demons and Santa's elves.

Belcher looked at monoclonal antibodies because they can be engineered to stick to many different things, but they proved hard to work with. Then, in the mid-1990s, she learned about the M13 phage, a long, skinny virus that parasitizes bacteria but is harmless to humans. The phage, which is roughly six nanometers wide and a micron long, encloses its single strand of DNA in a protein coat, with some 2,700 copies of one kind of protein lining the filamentous body and a few copies each of several other kinds of protein capping the ends. The different kinds of proteins can be engineered to vary from phage to phage, making for a billion possible permutations, each conferring particular chemical affinities. A phage might stick to one material along its sides, another at one end, and still another at the opposite end.

Such chemical specificity has long been exploited by biologists, who use M13 phages that bind to particular organic substances to identify unknown samples. Belcher was the first to demonstrate that the virus could also tag and manipulate inorganic molecules, such as the metals and semiconductors that lie at the heart of so many useful products. It was a rare example of the physical sciences borrowing from the biological sciences: because the biologists had done the spadework, Belcher could go out and buy a vast collection of phage variants, called a phage display library, for approximately $300.

To get a phage that binds to the right molecule, Belcher uses

3-D LIQUID CRYSTAL depicted in this artist's representation consists of multiple copies of a phage (a bacteria-infecting virus) called M13 (*gold*) that bound to inorganic nanocrystals (*pink*) and assembled themselves into an ordered array. Such films could be used in flexible displays.

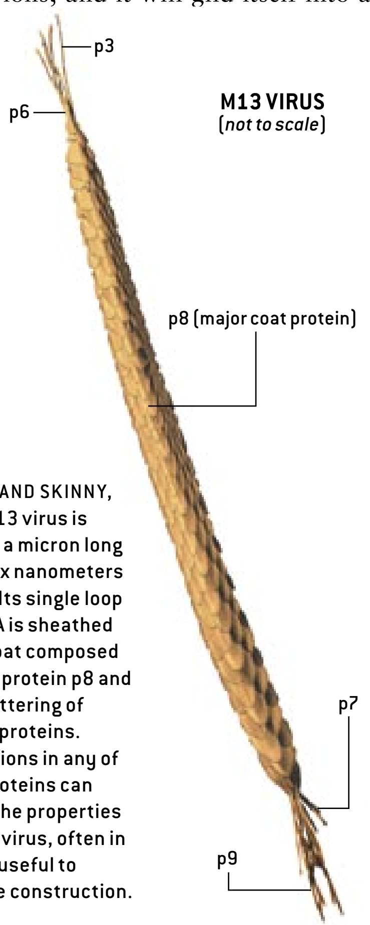
PHIL SAUNDERS *Space Channel Ltd.*

a process called directed evolution. "We throw our billion possibilities into a beaker with some material, wash it off and see what sticks to the material," she says. "We remove the ones that stick by changing their interaction with the surface, say, by lowering the pH, then collect whatever stuck and infect it into host bacteria, to amplify the phage."

Phage Feats

AMPLIFICATION provides trillions of copies of a promising subset of phages for another stage of evolution. This time the conditions in the solution are altered to make it a little harder for the phages to bond to the target material; again, the less sticky variants are washed off, the survivors are amplified, and the process is repeated under still more demanding conditions. At the end of the process, which can take three weeks, only one phage variant remains—the most selectively sticky of them all.

Put a phage with a highly specific taste for gold into a solution containing gold ions, and it will gild itself into a wirelet a micron long, suitable for connecting adjacent elements in a microcircuit. A variant of this phage will even link up with its fellows to form a gold wire many centimeters long, which can be spun like thread and woven into cloth fabric. Such a wire, bonded to chemically sensitive receptors, might detect toxic or biologically threatening agents.



LONG AND SKINNY, the M13 virus is about a micron long and six nanometers wide. Its single loop of DNA is sheathed in a coat composed of the protein p8 and a smattering of other proteins. Variations in any of the proteins can alter the properties of the virus, often in ways useful to device construction.

A year or two ago Belcher got yeast cells to fix gold, in an experiment that as yet has no practical applications (although the cells' six-micron width might make them easily visible as fluorescent markers in certain experiments). In the meantime, her M.I.T. students who are learning how to use organisms as a basis for materials perform the gold fixing as an exercise.

Although Belcher continues to examine the experimental merit of various other organisms, she focuses on M13, in part because its immense length-to-width ratio makes it naturally assemble into more complex shapes. "Think of crayons in a box," Belcher says. "If you shake just a few of them up, they'll settle randomly, but if you increase their concentration, they'll tend to pack." She has gotten selected M13 phages to form a film 10 centimeters square and less than a micron thick, a structure she then fixes into a stable sheet by throwing in the odd chemical cross-linkage.

Right now Belcher—along with M.I.T. collaborators Yet-Ming Chiang, Paula Hammond and Ki Tae Nam—is developing such sheets as electrodes for an ultralight lithium-ion battery, a project funded by the U.S. Army. "Battery weight is a big issue for them. The first few planes into Baghdad were full of batteries," Belcher says. "Our electrodes weigh 40 to 50 milligrams, versus grams for the conventional kind."

The negative electrode can be formed from a sheet of phages bred to encrust themselves in gold and cobalt oxide—the gold to increase conductivity, the cobalt oxide to exchange ions with the battery electrolyte. Such ion exchange is what moves charge from one electrode to the other. The electrode assembles directly on a prepatterned polymer electrolyte, forming a bilayer. Now the group is getting phages to grow a positive counterelectrode that will stick to the other side of the electrolyte.

A NEW KIND OF BATTERY

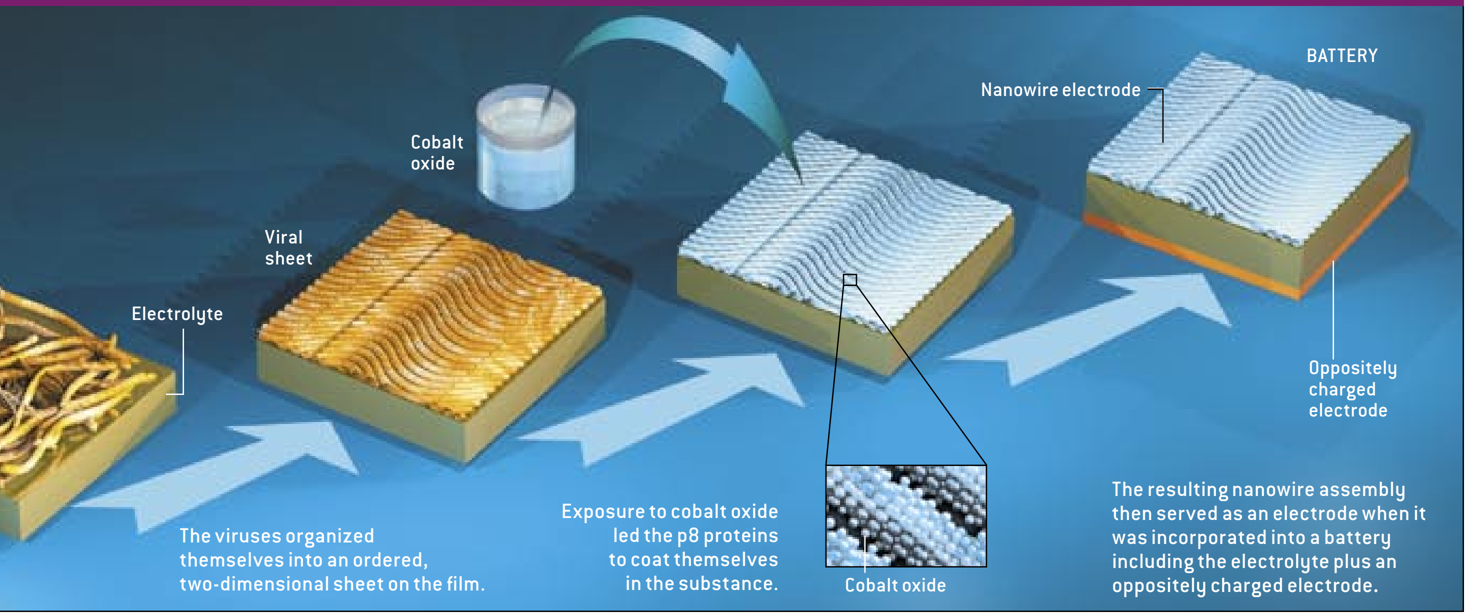
Angela Belcher and her colleagues at M.I.T. are testing selected M13 variants as building blocks for electrodes in flexible, lightweight lithium-ion batteries. One recent experiment is depicted.

The group selected an M13 variant that can bind to cobalt oxide, a substance able to store lithium ions. They put copies of the virus on an electrolyte film containing lithium.

The goal is to shape the sheets into a solid with positive and negative electrodes alternating on the surface so they may be connected in series for higher voltage. The short distances between electrodes permit fast charging and discharging and allow for the optimal use of components. The battery will also mold to whatever space a designer may have. It will thus save both weight and space, features critical to everything from military electronics to ultrathin MP3 players.

There seem to be no elements and compounds that phages cannot tell apart. One phage is specific for the semiconductor gallium arsenide and is insensitive to its close cousin gallium nitride, giving it a power of discrimination that might allow it to detect flaws in chips. Chipmakers sometimes grow crystals of one of these substances on top of some other semiconductor so that the slightly different spacing of the crystalline lattices will induce mechanical strain, which in turn will affect electronic behavior. When the crystals do not mesh properly, the occasional atom will jut out where it should not, creating a defect to which a phage can stick. If such a phage also bears a fluorescent tag, it will then glow under the right conditions, and a microscope can pinpoint the defect.

PHIL SAUNDERS *Space Channel Ltd.* (*left and above*)



The viruses organized themselves into an ordered, two-dimensional sheet on the film.

Exposure to cobalt oxide led the p8 proteins to coat themselves in the substance.

The resulting nanowire assembly then served as an electrode when it was incorporated into a battery including the electrolyte plus an oppositely charged electrode.

Big Plans

BELCHER WANTS TO TAKE the technology even further, though. "We want to see if we can transition over to finding manufacturing defects in things like an airplane wing," she says. Her group also wants to coax M13 phages into building complete transistors from molecules of semiconductors, and then churn them out by the billion. She admits that viral transistors might not be smaller and better, but because they would be made without harsh chemicals, their manufacture should yield less toxic waste.

Belcher also hopes to return a favor to biochemistry by getting M13 to bind both to cancer cells and to nanodevices known as quantum dots that show up in medical body scans. Quantum dots have yet to be tried in humans, in part because of concerns over the toxicity of their constituent heavy metals, notably cadmium. Belcher is trying to get her phage to attach to safer particles made of gallium nitride, indium nitride or some other semiconductor. The National Cancer Institute is funding this research.

Most of Belcher's M.I.T. projects are years away from finding commercial applications, but Cambrios needs to work on applications that can go to market within about two years, before it burns through its capital. Michael Knapp, president and CEO, notes that in the three years since it began, the company has raised $14 million in two rounds of financing, opened a lab and hired 20 people, which makes for a "burn rate" of $5 million a year. He says Cambrios is pursuing a niche product that will yield high profits on low volumes: a touch-sensitive screen on a flexible plastic sheet.

The army wants a flexible screen to slap onto a windshield so that a computer interface can be quickly put within the driver's field of vision; designers also want to build the screen into a computer display that can be rolled up when not in use, saving space. Today's production techniques cannot build flexible screens, because they work at temperatures that would melt the plastic backing.

"We're planning on launching a product in the middle of next year," Knapp says. "We'll be involved in the production one way or another, but because the electronics industry prefers to buy from someone they know, it's almost certain that we will have partners."

Belcher continues to consult regularly with Cambrios on its various projects, and she also conducts research on her own. She notes that although the company has rights to her viral manufacturing technique, she, along with M.I.T., retains the intellectual property from her current pursuits. Her M.I.T. group worked on the battery, for instance, whereas Cambrios devised the touch screen in-house.

"I really liked developing the basic science and transitioning to a company," she says, adding that she wants to do it again. She won't think out loud about what her next company will do, except to say that, like Cambrios, with its viral assemblers and inorganic building blocks, it will involve connecting things that do not usually go together. SA

Philip E. Ross is a science writer based in New York City.

MORE TO EXPLORE

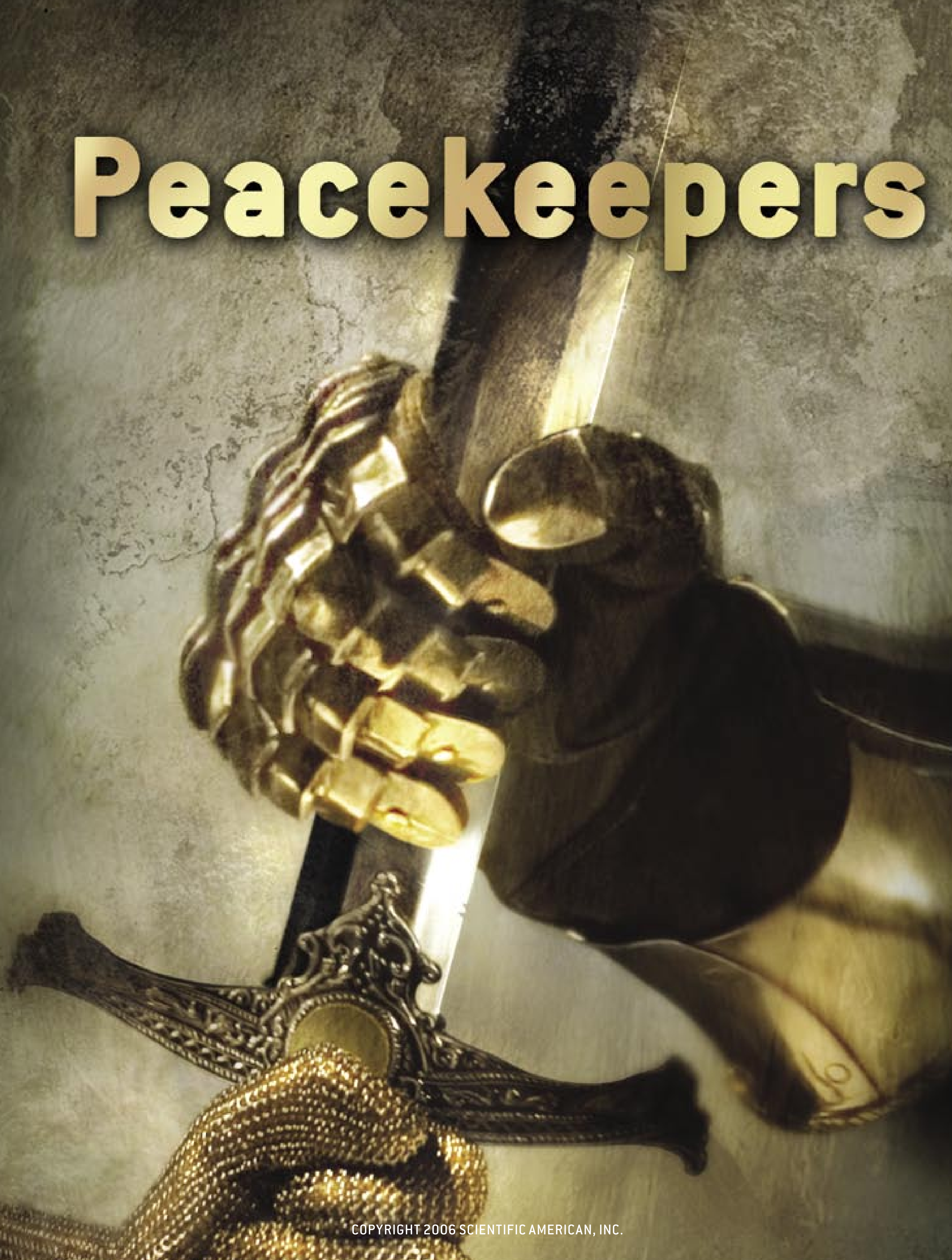
Ordering of Quantum Dots Using Genetically Engineered Viruses. Seung-Wuk Lee, Chuanbin Mao, Christine E. Flynn and Angela M. Belcher in *Science,* Vol. 296, pages 892–895; May 3, 2002.

Viruses as Vehicles for Growth, Organization and Assembly of Materials. Christine E. Flynn, Seung-Wuk Lee, Beau R. Peelle and Angela M. Belcher in *Acta Materialia,* Vol. 51, Issue 19, pages 5867–5880; November 25, 2003.

Spontaneous Assembly of Viruses on Multilayered Polymer Surfaces. Pil J. Yoo et al. in *Nature Materials,* Vol. 5, pages 234–240; March 2006.

Virus-Enabled Synthesis and Assembly of Nanowires for Lithium Ion Battery Electrodes. Ki Tae Nam et al. in *Science,* Vol. 312, pages 885–888; May 12, 2006.

Peacekeepers



of the IMMUNE SYSTEM

Regulatory T cells, only recently proven to exist, keep the body's defenses from attacking the body itself. Manipulations of these cells could offer new treatments for conditions ranging from diabetes to organ rejection

By Zoltan Fehervari and Shimon Sakaguchi

"Horror autotoxicus."

A century ago the visionary bacteriologist Paul Ehrlich aptly coined that term to describe an immune system attack against a person's own tissues. Ehrlich thought such autoimmunity—another term he coined—was biologically possible yet was somehow kept in check, but the medical community misconstrued his two-sided idea, believing instead that autoimmunity had to be inherently impossible. After all, what wrong turn of evolution would permit even the chance of horrendous, built-in self-destruction?

Slowly, though, a number of mysterious ailments came to be recognized as examples of horror autotoxicus—among them multiple sclerosis, insulin-dependent diabetes (the form that commonly strikes in youth) and rheumatoid arthritis. Investigators learned, too, that these diseases usually stem from the renegade actions of white blood cells known as CD4+ T lymphocytes (so named because they display a molecule called CD4 and mature in the thymus). Normal versions of these cells serve as officers in the immune system's armed forces, responsible for unleashing the system's combat troops against disease-causing microorganisms. But sometimes the cells turn against components of the body.

Ehrlich was correct in another way as well. Recent work has identified cells that apparently exist specifically to block aberrant immune behavior. Called regulatory T cells, they are a subpopulation of CD4+ T cells, and they are vital for maintaining an immune system in harmony with its host. Increasingly, immunologists are also realizing that these cells do much more than quash autoimmunity; they also influence the immune system's responses to infectious agents, cancer, organ transplants and pregnancy. We and others are working to understand exactly how these remarkable cells carry out their responsibilities and why they sometimes function imperfectly. The findings should reveal ways to regulate the regulators and thus to depress or enhance immune activity as needed and, in so doing, to better address some of today's foremost medical challenges.

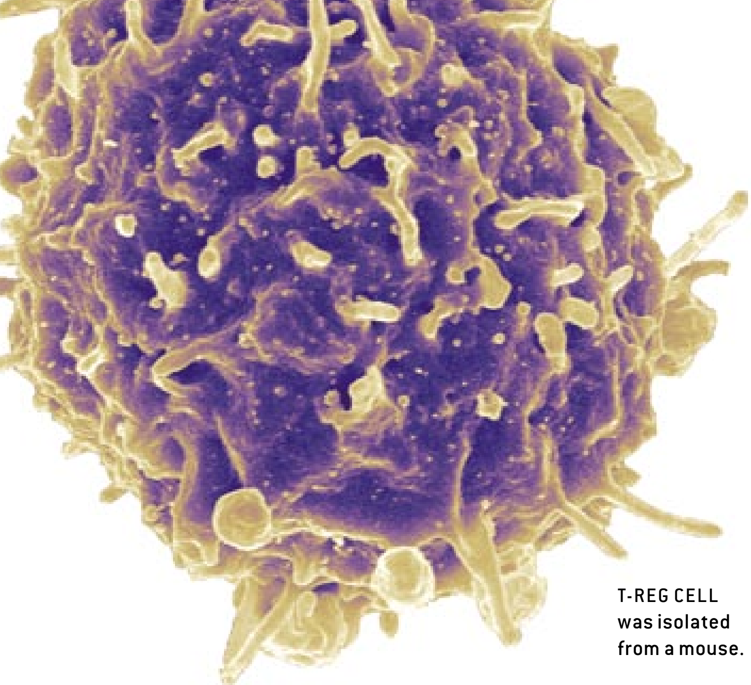
Imperfect Defenses

LIKE THE IMMUNOLOGISTS of Ehrlich's time, many people today would be dismayed to know that no matter how healthy they may be, their bodies harbor potentially destructive immune system cells quite capable of triggering autoimmune disease. Yet this immunological sword of Damocles can be easily demonstrated. If a mouse, for example, is injected with proteins from its own central nervous system, along with an adjuvant (a generalized immune system stimulus), a destructive immune reaction ensues. Much as in multiple sclerosis, T cells launch an attack on the animal's brain and spinal cord.

By varying the source of the injected self-protein, researchers can provoke other autoimmune diseases in laboratory animals—which indicates that potentially harmful immune sys-

WHEN THE IMMUNE SYSTEM wields its weapons inappropriately, regulatory T cells, also called T-regs, restrain them.

CLIFF NIELSEN



T-REG CELL was isolated from a mouse.

tem cells can mount self-attacks on a wide variety of tissues. The risk appears to hold true in humans, too, because autoreactive immune system cells can be captured readily from the blood of a healthy person. In a test tube, they react strongly to samples of that person's tissues.

Given such demonstrations of clear and imminent danger, investigators naturally wondered how it is that most animals and humans are untroubled by autoimmune disease. Put another way, they wanted to know how the immune system distinguishes threats such as microbes from a person's own tissues. They found that to achieve self-tolerance—the ability to refrain from attacking one's own organs—the immune system enlists numerous safeguards. The first defense, at least where T cells are concerned, occurs in the thymus, which lies inconspicuously in front of the heart. In the thymus, immature T cells undergo a strict "education" in which they are programmed to not react strongly (and therefore harmfully) to any bodily tissues. Disobedient cells are destroyed. No system is perfect, though, and in fact a small number of autoaggressive T cells slip through. Escaping into the bloodstream and into lymph vessels, they create the immune system's potential for unleashing autoimmune disease.

Blood and lymph vessels are where a second line of defense comes into play. This layer of protection against autoimmunity has several facets. Certain tissues, including those of the brain and spinal cord, are concealed from immune cell patrols simply by having a paucity of blood and lymph vessels that penetrate deep into the tissue. Their isolation, however, is not absolute, and at times, such as when the tissues are injured, self-reactive immune cells can find a way in. Additional modes of protection are more proactive. Immune cells showing an inappropriate interest in the body's own tissues can be targeted for destruction or rendered quiescent by other immune system components.

Among the immune cells that carry out these proactive roles, regulatory T cells may well be the most crucial. The majority, if not all of them, learn their "adult" roles within the thymus, as other T cells do, then go forth and persist throughout the body as a specialized T cell subpopulation.

Overview/*Immune Regulators*

- For years, immunologists doubted that cells specifically responsible for suppressing immune activity existed. But they do. They are called regulatory T cells.
- These so-called T-reg cells combat autoimmunity. They also help the body resist repeat infections by a returning invader, protect needed bacteria in the gut and aid in sustaining pregnancy. On the negative side, they abet cancer cells in escaping immune attacks.
- Ongoing research promises to yield new therapies for autoimmune disorders and cancer and could lead to treatments that would spare organ transplant recipients from having to take immunosuppressive drugs for life.

Discovering the Peacekeepers

FINDINGS HINTING AT the existence of regulatory T cells date back surprisingly far. In 1969 Yasuaki Nishizuka and Teruyo Sakakura, working at the Aichi Cancer Center Research Institute in Nagoya, Japan, showed that removing the thymus from newborn female mice had a curious outcome: the animals lost their ovaries. At first it was thought that the thymus must secrete some kind of hormone needed for survival of the developing ovaries. Later, though, it turned out that immune system cells invaded the ovaries. The ovarian destruction was therefore an autoimmune disease, which had presumably been unleashed by the animals' loss of a countervailing regulatory process. If the mice were inoculated with normal T cells, the autoimmune disease was inhibited. T cells, then, could at times police themselves somehow.

In the early 1970s John Penhale of the University of Edinburgh made analogous observations in adult rats, and Richard Gershon of Yale University became the first to propose the existence of a T cell population capable of damping immune responses, including autoaggressive ones. This hypothetical immune system member was christened the suppressor T cell. At the time, though, no researcher was able to actually find one or pinpoint the molecular action by which one immune system cell could restrain another. Consequently, the concept of the suppressor T cell languished along the fringes of mainstream immunology.

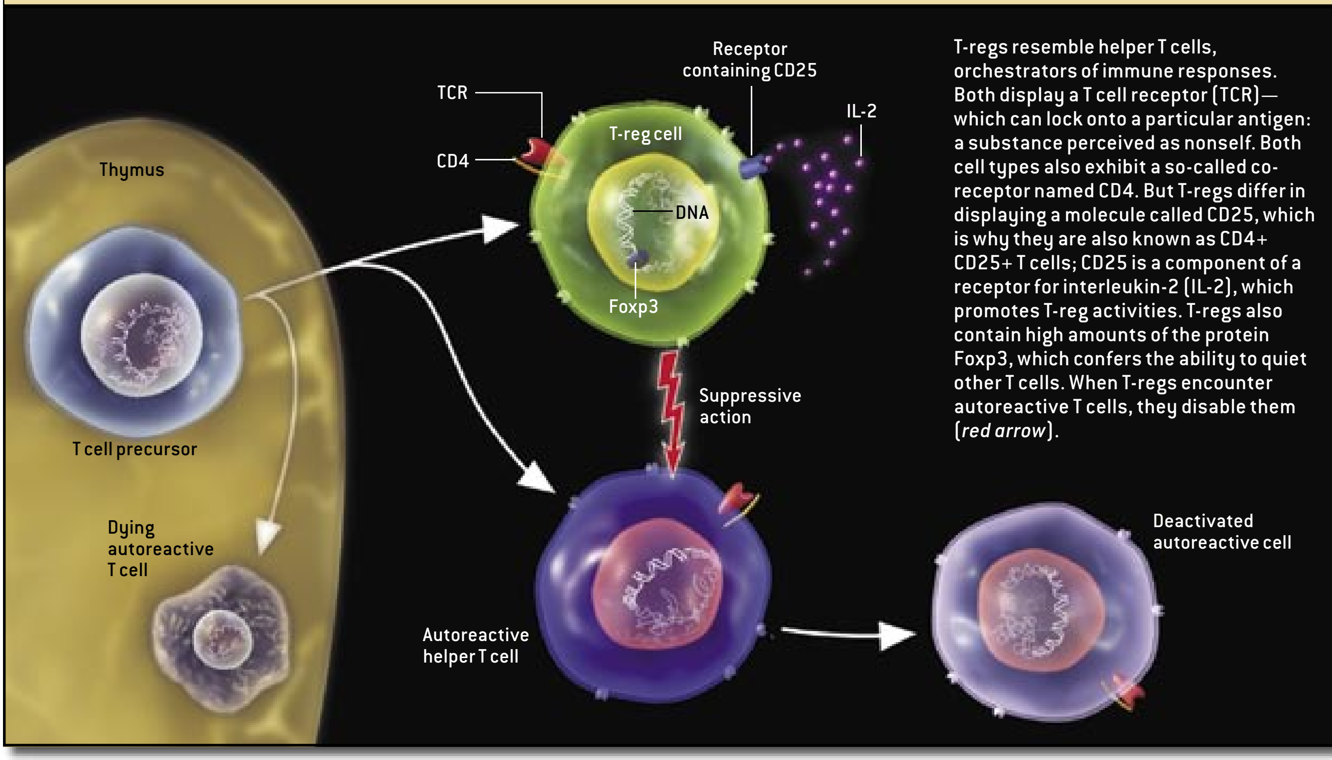
Despite the negative atmosphere, some researchers persisted in trying to identify T cells with an ability to prevent autoimmune disease. The basic hope was to discover a telltale molecular feature at the surface of such cells—a "marker" by which suppressor T cells could be distinguished from other cells. Beginning in the mid-1980s, various candidate markers were explored.

In 1995 one of us (Sakaguchi) finally demonstrated that a molecule called CD25 was a reliable marker. When, in studies of mice, he removed CD4+ T cells displaying that molecule, organs such as the thyroid, stomach, gonads, pancreas and salivary glands came under an autoimmune attack character-

KIM J. HASENKRUG AND DAVID W. DORWARD

MECHANISMS OF TOLERANCE

T-reg cells help to ensure that immune system components—including T cells that fight infections—refrain from attacking normal tissues. The thymus, where all T cell varieties mature, directly eliminates many strongly autoreactive cells (*left*), but its vigilance is imperfect, so T-regs patrol the body in search of renegades (*right*).



T-regs resemble helper T cells, orchestrators of immune responses. Both display a T cell receptor (TCR)—which can lock onto a particular antigen: a substance perceived as nonself. Both cell types also exhibit a so-called co-receptor named CD4. But T-regs differ in displaying a molecule called CD25, which is why they are also known as CD4+ CD25+ T cells; CD25 is a component of a receptor for interleukin-2 (IL-2), which promotes T-reg activities. T-regs also contain high amounts of the protein Foxp3, which confers the ability to quiet other T cells. When T-regs encounter autoreactive T cells, they disable them (*red arrow*).

ized by dramatic inflammation: white blood cells swarmed into the organs and damaged them.

In an important confirmatory experiment, T cell populations obtained from normal mice were depleted of their CD4+ CD25+ T cells, which evidently made up only a small proportion (at most, 10 percent) of the overall T cell pool. Then T cells left behind were transferred to mice engineered to lack an immune system of their own. This maneuver caused autoimmune disease. And the more complete the depletion was in the donor animals, the more severe the spectrum of disease became in the recipients—with comprehensive depletion often proving to be fatal. Reintroducing CD4+ CD25+ T cells, even in small numbers, conferred normal immunity and protected the animals from these disorders. Experiments conducted wholly in test tubes also produced valuable confirmatory evidence. Perhaps to absolve "suppressor cells" of any lingering stigma, immunologists started to call them CD25+ regulatory T cells, or simply T-regs.

How Do T-regs Work?

TO THIS DAY, the precise ways in which T-regs suppress autoimmune activity have remained mysterious, making their function a continuing subject of intense inquiry. The cells appear capable of suppressing a wide variety of immune system cells, impeding the cells' multiplication and also their other activities, such as secretion of cell-to-cell chemical signals (cytokines). And researchers tend to agree that T-regs are activated by direct cell-to-cell contacts. Beyond that, the picture is rather murky [*see box on next two pages*].

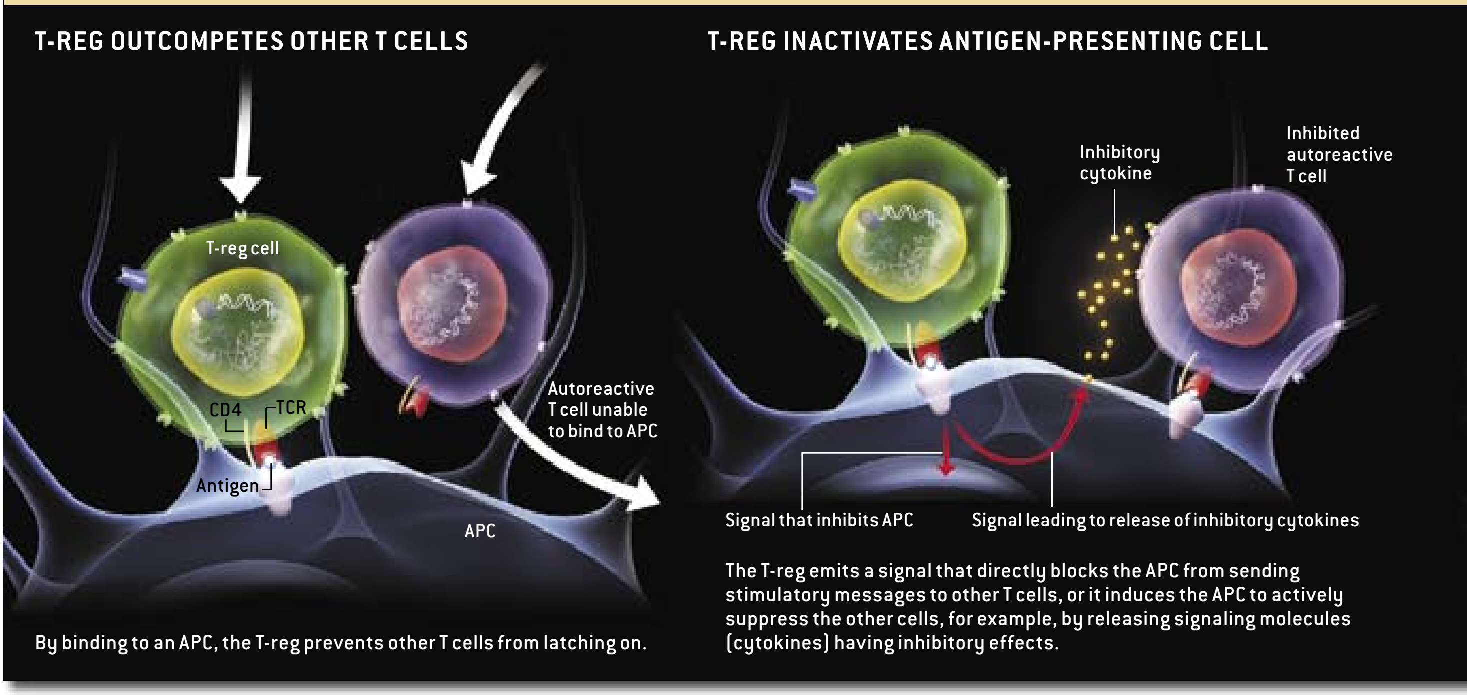
Recently, however, our laboratory at Kyoto University and, independently, Alexander Rudensky's group at the University of Washington and Fred Ramsdell's group at CellTech R&D in Bothell, Wash., found a fresh clue as to how T-regs develop and function. The cells contain a large amount of an intracellular molecule called Foxp3. In fact, the enrichment is greater than has been reported for any other T-reg molecular feature.

Foxp3 is a transcription factor: a molecule that regulates the activity of specific genes, thereby controlling a cell's production of the protein that each such gene encodes. Because proteins are the main worker molecules in cells, altered production of one or more of them can affect how a cell functions. In the case of Foxp3, the changes it induces in gene activity apparently turn developing T cells into T-regs. Indeed, artificially introducing Foxp3 into otherwise unremarkable T cells provokes a reprogramming, by which the cells acquire all the suppressive abilities of full-fledged T-regs produced by the thymus. A type of mouse called the Scurfy strain, long known to researchers, has recently been found to have only an inactive, mutant form of the Foxp3 protein, along with a total absence of T-regs. The consequence is an immune system gone haywire, with massive inflammation

AMADEO BACHAR

HOW DO T-REGS PREVENT AUTOIMMUNITY?

No one fully understands how T-regs block autoimmune attacks. Three reasonable possibilities follow. All three involve interfering with a key step in triggering immune responses: signaling between T cells and antigen-presenting cells (APCs). Before helper T cells will call forth other troops and before "cytotoxic" T cells will attack tissue perceived to be infected, APCs must display antigens for the cells' perusal. If the T cell receptor (TCR) of a helper or cytotoxic cell recognizes a displayed antigen and also receives certain other signals from the APC, the T cells will



T-REG OUTCOMPETES OTHER T CELLS

By binding to an APC, the T-reg prevents other T cells from latching on.

T-REG INACTIVATES ANTIGEN-PRESENTING CELL

The T-reg emits a signal that directly blocks the APC from sending stimulatory messages to other T cells, or it induces the APC to actively suppress the other cells, for example, by releasing signaling molecules (cytokines) having inhibitory effects.

AMADEO BACHAR

in numerous organs, leading to the animals' early death.

Of course, investigators study T-regs in animals such as mice so that the knowledge gained may be applied to humans. So what evidence is there that T-regs are indeed important in humans—or that they exist in us at all?

It turns out that the molecular features characteristic of T-regs in rodents are also characteristic of a subset of T cells in humans. In humans, as in rodents, these cells exhibit the CD25 molecule and have a high content of Foxp3. In addition, the cells are immunosuppressive, at least in a test tube.

Perhaps the most compelling indications that they are vital to human health come from a rare genetic abnormality called IPEX (*i*mmune dysregulation, *p*olyendocrinopathy, *e*nteropathy, *X*-linked syndrome). Arising from mutations in a gene on the X chromosome, IPEX affects male children, who unlike females inherit only one X chromosome and hence have no chance of inheriting a second, normal copy of the gene, which would encode a healthy version of the affected protein. In males the mutation results in autoimmune disease affecting multiple organs, including the thyroid and (as happens in insulin-dependent diabetes) the pancreas, and also in chronic intestinal inflammation (inflammatory bowel disease) and uncontrolled allergy (food allergy and severe skin inflammation), all of which can be understood as varied manifestations of the hyperactivity of an immune system unrestrained by T-regs. Death comes in infancy or soon after, with contributing causes ranging from autoimmune diabetes to severe diarrhea. The specific genetic flaw underlying IPEX has recently proved to be mutation in none other than *Foxp3*. IPEX is therefore the human counterpart of the illness in Scurfy mice.

Beyond Self-Tolerance

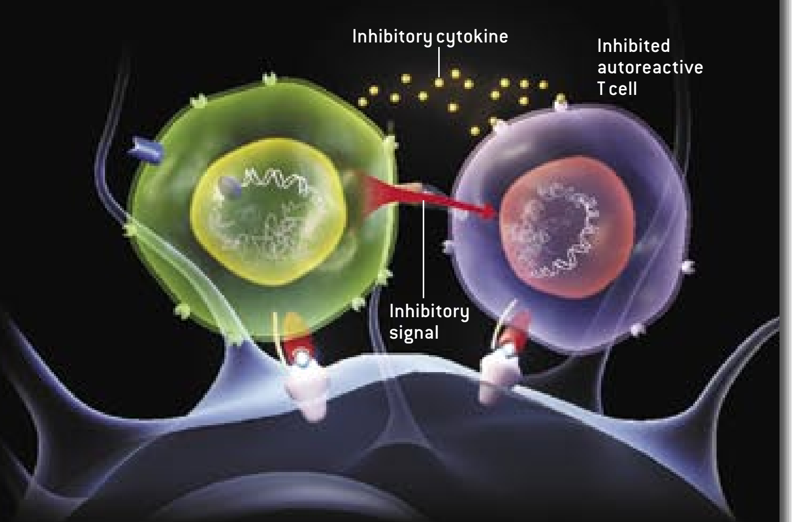
THE EVIDENCE, THEN, indicates that T-regs do prevent autoimmune disease in humans. But the cells also appear to serve health in other ways, including participating (in some surprising ways) in responses to microbes.

Throughout the 1990s Fiona Powrie and her colleagues at the DNAX Research Institute in Palo Alto, Calif., experimented with transferring T cell populations depleted of T-regs into mice engineered to lack an immune system of their own. In one set of studies, the transfer induced a severe, often fatal form of inflammatory bowel disease. But the aberrant immune activity was not directed primarily at bowel tissue itself.

The bowels of rodents, like those of humans, are home to a vast bacterial population, typically more than a trillion for every gram of intestinal tissue. Although these bacteria are foreign, they are usually far from harmful; indeed, they promote the digestion of food and even displace dangerous bacteria, such as salmonella, that would otherwise try to colonize the intestines. Normally the immune system tolerates the presence of the helpful population. But in Powrie's mice, it attacked. And in doing so, the transplanted immune cells caused collateral damage to the recipient's gut. Yet transfer of T-regs caused no problems. In fact, if the T-regs were transferred

become active against the bearer of that antigen—even if the antigen is from the body itself, instead of from an infectious agent. The TCRs of T-regs also recognize particular antigens, and they specifically suppress T cells that focus on those same antigens.

T-REG QUIETS OTHER T CELLS DIRECTLY



The T-reg uses the APC essentially as a platform for stabilizing contact with another T cell bound to the APC. Then the T-reg sends an inhibitory signal directly into the T cell or emits inhibitory molecules that act at close range.

along with the other T cells, they prevented the bowel disease that would otherwise have ensued. Overall, the immune system appeared to be on a hair trigger, prepared to assault gut bacteria and held in check only by T-regs.

A similar hair trigger may affect the immune system's responses to harmful foreigners. On the one hand, T-regs might rein in an overemphatic response. On the other hand, the reining in might keep an invader from being totally destroyed, enabling it to persist and potentially flare up again. For example, some findings suggest that failure to clear the stomach of a bacterium called *Helicobacter pylori,* now known to cause stomach ulcer, stems from blunting by T-regs of the immune system's weaponry.

Work by David Sacks and his colleagues at the National Institutes of Health has revealed further complexity. It implies that leaving a few survivors among invading organisms may not be entirely a bad thing. The researchers infected mice with a fairly innocuous parasite. Even when the immune system was fully intact, it allowed a small number of parasites to remain, after which reinfection triggered a prompt, efficient response. If the immune system was depleted of its T-regs, however, the parasite was completely purged, but reinfection was dealt with inefficiently, as if the mice had never before encountered the invader. Hence, T-regs appear to contribute to maintaining immunological memory, a process that is crucial for immunity to repeated infection and that also underlies the success of vaccination.

Research hints, too, at a role for T-regs in protecting pregnancies. Every pregnancy unavoidably poses quite a challenge to the mother's immune defenses. Because the fetus inherits half its genes from the father, it is genetically half-distinct from its mother and thus is in essence an organ transplant. Within the trophoblast, the placental tissue that attaches the fetus to the uterine wall, a number of mechanisms give the fetus some safety from what would amount to transplant rejection. The trophoblast not only presents a physical barrier to would-be attackers in the mother's blood but also produces immunosuppressive molecules.

The mother's immune system seems to undergo changes as well. Reports of women in whom an autoimmune disease such as multiple sclerosis abates during pregnancy provide anecdotal evidence that T-regs become more active. Some recent experiments offer more direct support. At the University of Cambridge, Alexander Betz and his colleagues have shown that during pregnancy in mice, maternal T-regs expand in number. Conversely, an experimentally engineered absence of T-regs leads to fetal rejection marked by a massive infiltration of immune cells across the maternal-fetal boundary. It is tempting to speculate that in some women, insufficient T-reg activity may underlie recurrences of spontaneous abortion.

Recruiting the Regulators

IN T-REGS, nature clearly has crafted a potent means of controlling immune responses. Tapping into this control would make T-regs a potentially powerful therapeutic ally against a wide range of medical disorders. It is still too early to expect to see applications in doctors' offices, but the available data suggest that delivering T-regs themselves, or perhaps medicines that increase or decrease their activity, could provide novel treatments for a variety of conditions. Indeed, some human trials are under way.

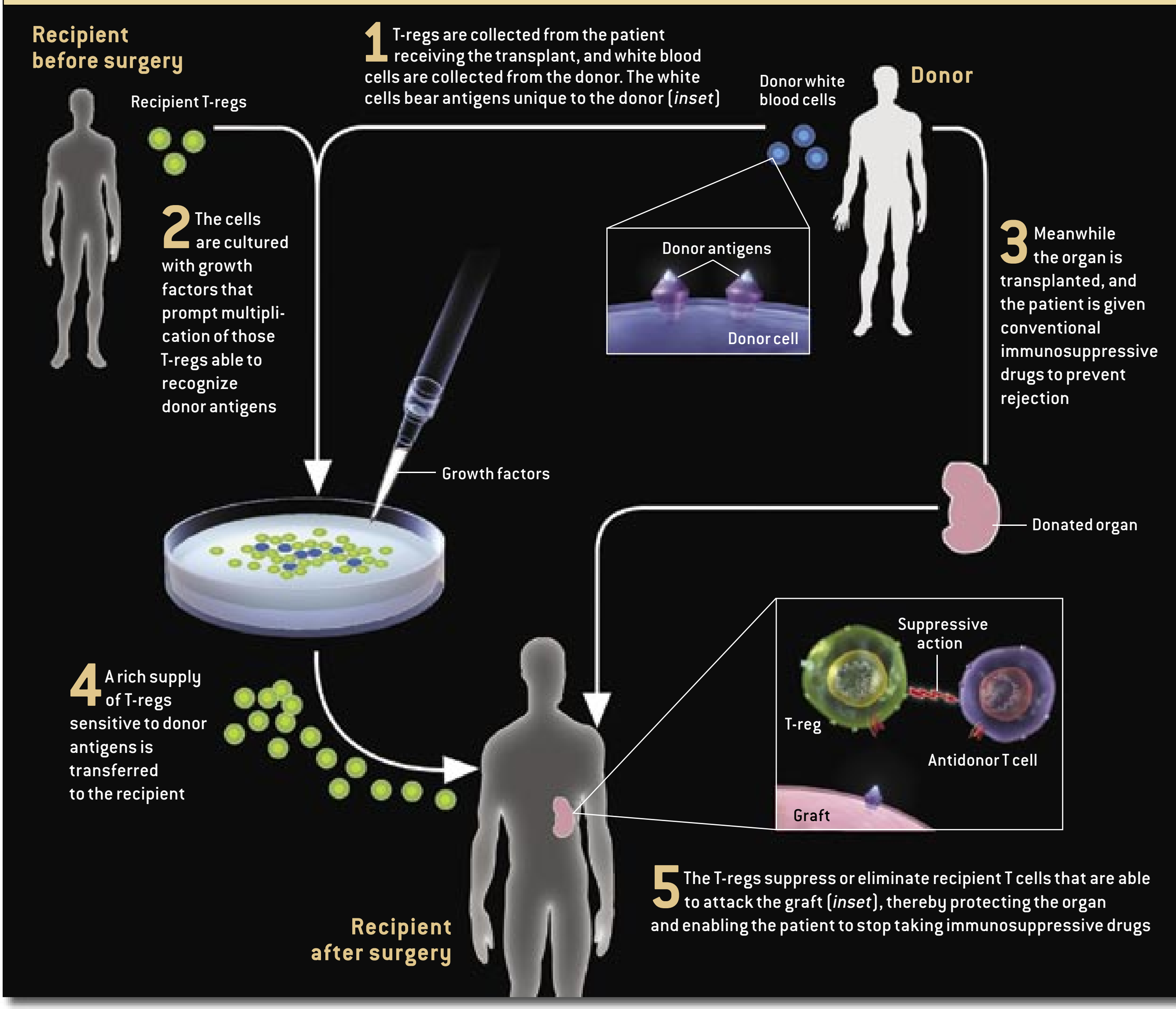
The most obvious application would involve enhancing T-reg activity to fight autoimmune diseases, and drug therapy is being explored in patients with multiple sclerosis and psoriasis, among other conditions. Pumping up T-reg activity might also be useful for treating allergies. The ease with which T-regs can keep immune responses at bay suggests that T-reg-based therapies could hold particular promise for preventing rejection of transplanted organs. The ideal would be for transplant recipients to tolerate grafts as well as they do their own tissues. Also ideal would be a tolerance that

THE AUTHORS

ZOLTAN FEHERVARI and *SHIMON SAKAGUCHI* began collaborating in 2002, when Fehervari took a postdoctoral position in Sakaguchi's laboratory at the Institute for Frontier Medical Sciences of Kyoto University in Japan. Fehervari is now a research associate in the department of pathology at the University of Cambridge, where he earned a Ph.D. in immunology. Sakaguchi is professor and chair of the department of experimental pathology at Kyoto. He began searching for regulatory T cells in the early 1980s and has studied them ever since.

ORGAN TRANSPLANTATION OF TOMORROW?

One day T-reg-based therapy may help preserve transplanted organs while limiting the amount of time a patient has to take immunosuppressive drugs, which can have undesirable side effects. A protocol might look like the following:



AMADEO BACHAR

endures as a permanent state of affairs, without need for immunosuppressive drugs, which can have many side effects.

The opposite type of T-reg-based therapy would be a selective depletion of T-regs to counter unwanted immunosuppression and, consequently, to strengthen beneficial immune responses. In practice, a partial depletion might be preferred to a complete one, because it should pose less risk of inducing autoimmune disease. Best of all would be removal solely of those T-regs that were specifically blocking a useful immune response. The depletion strategy might be especially advantageous against infectious diseases that the immune system, left to itself, tends to combat inadequately—perhaps tuberculosis or even AIDS.

In addition, T-reg reduction might be advantageous for fighting cancer. Much evidence suggests that circulating immune cells keep a lookout for molecular aberrations that occur as a cell becomes cancerous. To the extent that T-regs impede this surveillance, they might inadvertently help a malignancy take root and grow. In fact, some cancers appear to encourage such help: they secrete molecular signals capable of attracting T-regs and of converting non-T-regs into T-regs. Some findings suggest, for example, that cancer patients have abnormally high numbers of active T-regs both in their blood and in the tumors themselves. Much of today's research into therapeutic manipulations of T-regs focuses on cancer.

Technical Challenges

SO FAR INVESTIGATORS are finding it challenging to develop medicines able to deplete or expand T-reg populations within a patient's body. To be most useful, these drugs would

Some T-Reg-Based Therapies under Study

The therapies listed below are among those in or likely to enter human trials. Most of the drugs under study aim to deplete or inhibit T-reg cells, so as to increase antitumor immune responses normally tempered by the cells. Delivery of such agents into the body would need to be managed carefully, though, to ensure that reducing T-reg activity does not lead to autoimmunity.

EFFECT ON T-REGS	EXAMPLES OF DISORDERS BEING TARGETED	TREATMENT APPROACHES
Depletion or inhibition (to enhance immunity)	Cancers of the skin (melanoma), ovary, kidney	A toxin fused to a substance, such as interleukin-2, able to deliver the toxin to T-regs Monoclonal antibodies (which bind to specific molecules) that have shown an ability to induce T-reg death or to block the cells' migration into tumors
Multiplication in patient (to dampen autoimmunity)	Multiple sclerosis, psoriasis, Crohn's disease, insulin-dependent diabetes	Vaccine composed of T cell receptor constituents thought to stimulate T-reg proliferation A monoclonal antibody that appears to stimulate T-regs by binding to a molecule called CD3
Multiplication in the lab, for delivery to patient	Graft-versus-host disease (immune cells in donated bone marrow attack recipient tissue)	Culture donor T-regs with selected antibodies and growth factors, then deliver resulting T-reg population before or at the time of bone marrow transplant (for prevention) or if graft-versus-host disease arises

usually need to act on the subsets of T-regs that have roles in a particular disorder, yet scientists often do not know precisely which T-regs to target.

Devising therapies based on administering T-regs themselves is difficult as well. One of the main obstacles is the need to obtain enough of the cells. Although researchers have found that T-regs can operate at low abundance relative to the cells they are suppressing, control of a human autoimmune disease would probably require tens of millions of T-regs. Acquiring such numbers of these relatively rare cells from a person's circulation might be impossible. Accordingly, some technique to expand their numbers outside the body would seem to be imperative.

Luckily, it also seems that this numbers game can be won. Worldwide, several research groups have reported that cells with immunosuppressive actions can be generated in relatively large numbers by treating ordinary T cells with a well-defined "cocktail" of biochemical signals. Whether the engendered cells, termed Tr1 cells, are identical to T-regs remains unclear, but it is beyond dispute that the cells are profoundly immunosuppressive.

Now that Foxp3 is known to be a key molecule controlling the development and function of T-regs, investigators may also be able to tailor-make large numbers of regulatory cells by using fairly standard laboratory techniques to transfer the *Foxp3* gene into more prevalent, and thus more easily obtainable, types of T cells. We and others are pursuing this approach intently and are also trying to identify the molecular events that switch on Foxp3 production during T-reg development. This knowledge might enable pharmaceutical researchers to fashion drugs specifically for that purpose, so that processing of cells outside the body and then infusing them would not be necessary.

For organ transplant patients, another way to obtain useful T-regs is under consideration. The procedure would involve removing T-regs from a prospective transplant recipient and culturing them with cells from the organ donor in a way that causes the T-regs most capable of suppressing rejection to multiply [*see box on opposite page*]. In rodents, T-regs generated in this manner have worked well. One of us (Sakaguchi) has shown, for example, that injection of a single dose of such T-regs at the time of skin grafting results in the graft's permanent acceptance, even though transplanted skin typically is rejected strongly. Meanwhile the treatment left the rest of the immune system intact and ready to fend off microbial invaders. The abundant research into T-regs suggests that such an approach can become a reality for humans and could be used to protect new transplant recipients until medications able to produce the same benefit more simply are developed.

Over the past decade, researchers' understanding of the immune system and how it governs its own actions has changed profoundly. In particular, it is now recognized that although the system permits potentially autodestructive T cells to circulate, it also deploys T cells capable of controlling them. Knowledge of how they develop and how they perform their remarkable immunosuppressive activities will be key in recruiting them for use against a host of debilitating and even fatal disorders. In permitting destruction of nonself while preventing destruction of self, T-regs may prove to be the ultimate immunological peacekeepers. SA

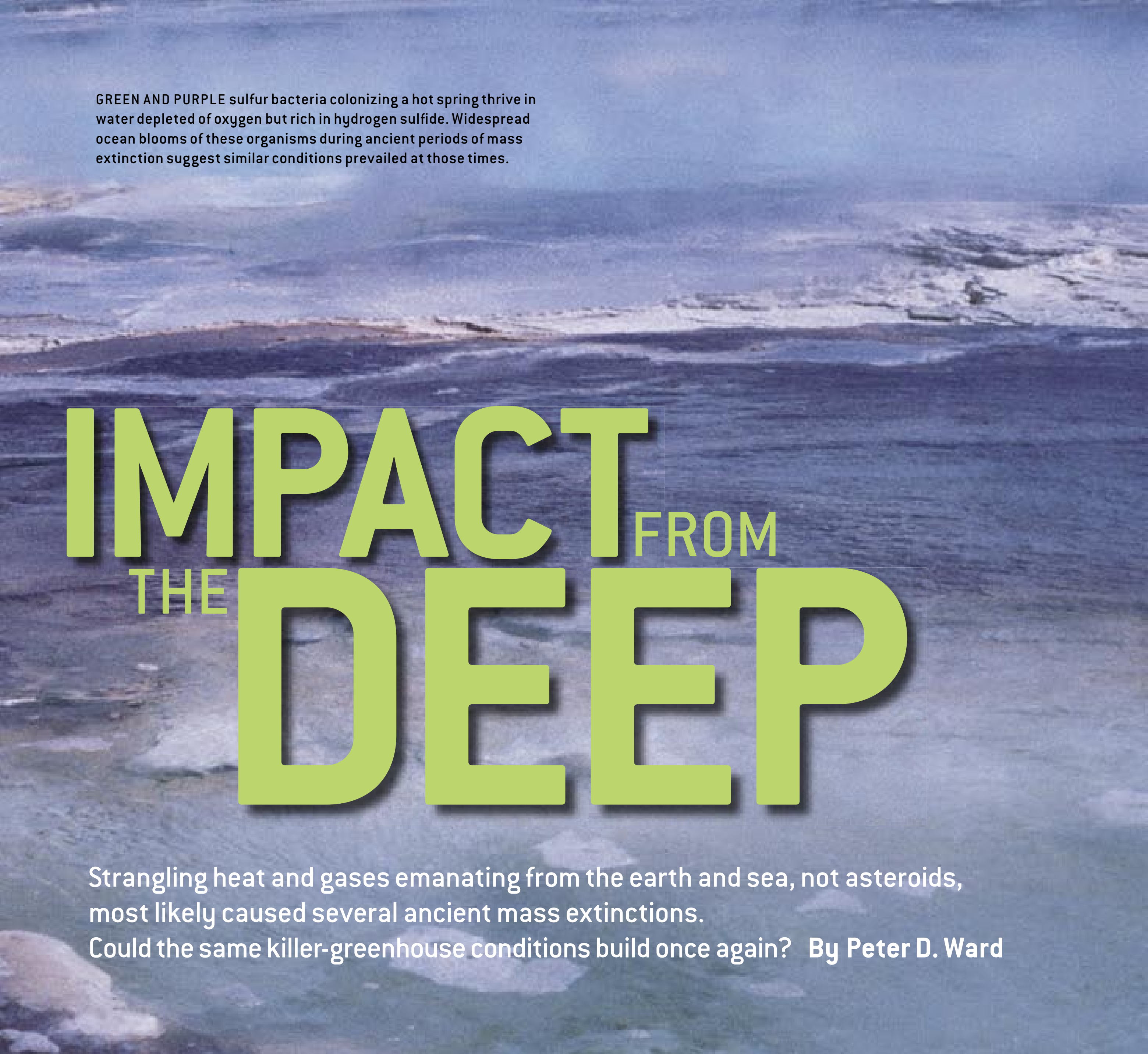
MORE TO EXPLORE

Naturally Arising CD4+ Regulatory T Cells for Immunologic Self-Tolerance and Negative Control of Immune Responses. Shimon Sakaguchi in *Annual Review of Immunology*, Vol. 22, pages 531–562; 2004.

Regulatory T-Cell Therapy: Is It Ready for the Clinic? J. A. Bluestone in *Nature Reviews Immunology*, Vol. 5, No. 4, pages 343–349; April 2005.

Regulatory T Cells, Tumour Immunity and Immunotherapy. Weiping Zou in *Nature Reviews Immunology*, Vol. 6, No. 4, pages 295–307; April 2006.

T Lymphocytes: Regulatory. Zoltan Fehervari and Shimon Sakaguchi in *Encyclopedia of Life Sciences*. Wiley InterScience, 2006. Available at **www.els.net**



GREEN AND PURPLE sulfur bacteria colonizing a hot spring thrive in water depleted of oxygen but rich in hydrogen sulfide. Widespread ocean blooms of these organisms during ancient periods of mass extinction suggest similar conditions prevailed at those times.

IMPACT FROM THE DEEP

Strangling heat and gases emanating from the earth and sea, not asteroids, most likely caused several ancient mass extinctions. Could the same killer-greenhouse conditions build once again? By Peter D. Ward

Philosopher and historian Thomas S. Kuhn has suggested that scientific disciplines act a lot like living organisms: instead of evolving slowly but continuously, they enjoy long stretches of stability punctuated by infrequent revolutions with the appearance of a new species—or in the case of science, a new theory. This description is particularly apt for my own area of study, the causes and consequences of mass extinctions—those periodic biological upheavals when a large proportion of the planet's living creatures died off and afterward nothing was ever the same again.

Since first recognizing these historical mass extinctions more than two centuries ago, paleontologists believed them to have been gradual events, caused by some combination of climate change and biological forces such as predation, competition and disease. But in 1980 the understanding of mass extinctions underwent a Kuhnian revolution when a team at the University of California, Berkeley, led by geologist Walter Alvarez proposed that the famous dinosaur-killing extinction 65 million years ago occurred swiftly, in the ecosystem catastrophe that followed an asteroid collision. Over the ensuing two decades, the idea that a bolide from space could smite a significant segment of life on the earth was widely embraced—and many researchers eventually came to believe that cosmic detritus probably caused at least three more of the five largest mass extinctions. Public acceptance of the notion crystallized with Hollywood blockbusters such as *Deep Impact* and *Armageddon*.

Now still another transformation in our thinking about life's punctuated past is brewing. New geochemical evidence is coming from the bands of stratified rock that delineate mass

extinction events in the geologic record, including the exciting discovery of chemical residues, called organic biomarkers, produced by tiny life-forms that typically do not leave fossils. Together these data make it clear that cataclysmic impact as a cause of mass extinction was the exception, not the rule. In most cases, the earth itself appears to have become life's worst enemy in a previously unimagined way. And current human activities may be putting the biosphere at risk once again.

After Alvarez

TO UNDERSTAND the general enthusiasm for the impact paradigm, it helps to review the evidence that fueled it. The scenario advanced by Alvarez, along with his father, physicist Luis W. Alvarez, and nuclear chemists Helen V. Michel and Frank Asaro, contained two separate hypotheses: first, that a fairly large asteroid—estimated to have been 10 kilometers in diameter—struck the earth 65 million years ago; second, that the environmental consequences of the impact snuffed out more than half of all species. They had found traces left by the blow in a thick layer of iridium—rare on the earth but common in extraterrestrial materials—that had dusted the globe.

The earth can, and probably did, exterminate its own.

Within a decade of this prodigious announcement the killer's thumbprint turned up, in the form of the Chicxulub crater hiding in plain sight on the Yucatán Peninsula of Mexico. Its discovery swept aside most lingering doubts about whether the reign of the dinosaurs had ended with a bang. At the same time, it raised new questions about other mass extinction events: If one was caused by impact, what about the rest? Five times in the past 500 million years most of the world's life-forms have simply ceased to exist. The first such event happened at the end of the Ordovician period, some 443 million years ago. The second, 374 million years ago, was near the close of the Devonian. The biggest of them all, the Great Dying, at the end of the Permian 251 million years ago, wiped out 90 percent of ocean dwellers and 70 percent of plants, animals, even insects, on land [see "The Mother of Mass Extinctions," by Douglas H. Erwin; SCIENTIFIC AMERICAN, July 1996]. Worldwide death happened again 201 million years ago, ending the Triassic period, and the last major extinction, 65 million years ago, concluded the Cretaceous with the aforementioned big bang.

In the early 1990s paleontologist David Raup's book *Extinctions: Bad Genes or Bad Luck?* predicted that impacts ultimately would be found to be the blame for all these major mass extinctions and other, less severe events as well. Evidence for impact from the geologic boundary between the Cretaceous and Tertiary (K/T) periods certainly was and remains convincing: in addition to the Chicxulub crater and the clear iridium layer, impact debris, including pressure-shocked stone scattered across the globe, attests to the blow. Further chemical clues in ancient sediments document rapid changes in the world's atmospheric composition and climate that soon followed.

For several other extinction periods, the signs also seemed to point "up." Geologists had already associated a thin iridium layer with the end Devonian extinctions in the early 1970s. And by 2002 separate discoveries suggested impacts at the end Triassic and end Permian boundaries. Faint traces of iridium registered in the Triassic layer. And for the Permian, distinctive carbon "buckyball" molecules believed to contain trapped extraterrestrial gases added another intriguing clue [see "Repeated Blows," by Luann Becker; SCIENTIFIC AMERICAN, March 2002]. Thus, many scientists came to suspect that asteroids or comets were the source of four of the "big five" mass extinctions; the exception, the end Ordovician event, was judged the result of radiation from a star exploding in our cosmic neighborhood.

As researchers continued to probe the data in recent years, however, they found that some things did not add up. New fossil analyses indicated that the Permian and Triassic extinctions were drawn-out processes spanning hundreds of thousands of years. And newly obtained evidence of the rise and fall of atmospheric carbon, known as carbon cycling, also seemed to suggest that the biosphere suffered a long-running series of environmental insults rather than a single, catastrophic strike.

Not So Sudden Impact

THE LESSON of the K/T event was that a large-body impact is like a major earthquake leveling a city: the disaster is sudden, devastating, but short-lived—and after it is over, the city quickly begins rebuilding. This tempo of destruction and subsequent recovery is reflected in carbon-isotope data for the K/T extinctions as well as in the fossil record, although verifying the latter took the scientific community some time. The expected sudden die-off at the K/T boundary itself was indeed visible among the smallest and most numerous fossils, those of the calcareous and siliceous plankton, and in the

Overview/*Mass Extinctions*

- More than half of all life on the earth has been wiped out, repeatedly, in mass extinctions over the past 500 million years.
- One such disaster, which included the dinosaurs' disappearance, is widely attributed to an asteroid impact, but others remain inadequately explained.
- New fossil and geochemical evidence points to a shocking environmental mechanism for the largest of the ancient mass extinctions and possibly several more: an oxygen-depleted ocean spewing poisonous gas as a result of global warming.

NATIONAL PARK SERVICE (*preceding pages*)

spores of plants. But the larger the fossils in a group, the more gradual their extinction looked.

Slowly, paleontologists came to understand that this apparent pattern was influenced by the sparsity of large-fossil samples for most of the soil and rock strata being studied. To address this sampling problem and gain a clearer picture of the pace of extinction, Harvard University paleontologist Charles Marshall developed a new statistical protocol for analyzing ranges of fossils. By determining the probability that a particular species has gone extinct within a given time period, this analytical method teases out the maximum amount of information yielded by even rare fossils.

In 1996 Marshall and I joined forces to test his system on K/T stratigraphic sections and ultimately showed that what had appeared to be a gradual extinction of the most abundant of the larger marine animals, the ammonites (molluscan fossils related to the chambered nautilus) in Europe, was instead consistent with their sudden disappearance at the K/T boundary itself. But when several researchers, including myself, applied the new methodology to earlier extinctions, the results differed from the K/T sections. Studies by my group of strata representing both marine and nonmarine environments during the latest parts of the Permian and Triassic periods showed a more gradual succession of extinctions clustered around the boundaries.

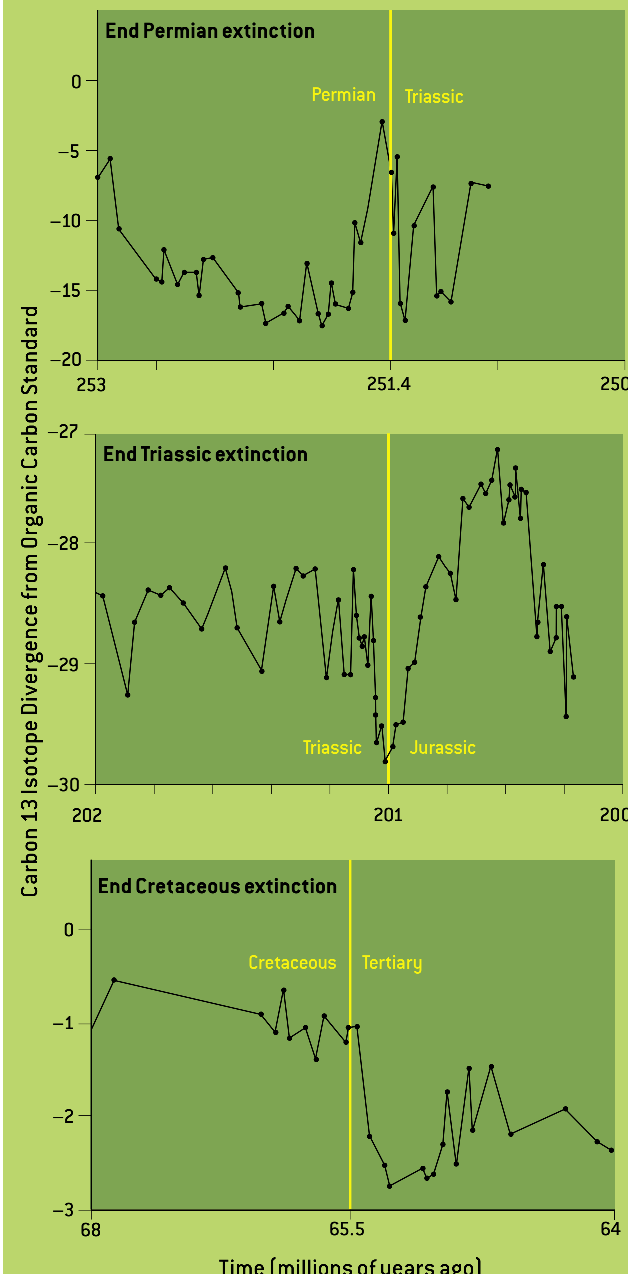
That pattern was also mirrored in the carbon-isotope record, which is another powerful tool for understanding rates of extinction. Carbon atoms come in three sizes, or isotopes, with slightly varying numbers of neutrally charged particles in the nucleus. Many people are familiar with one of these isotopes, carbon 14 (^{14}C), because its decay is often used to date specific fossil skeletons or samples of ancient sediments. But for interpreting mass extinctions, a more useful type of information to extract from the geologic record is the ratio of ^{12}C to ^{13}C isotopes, which provides a broader snapshot of the vitality of plant life at the time.

That is because photosynthesis largely drives changes in the ^{12}C-^{13}C ratio. Plants use energy from the sun to split carbon dioxide (CO_2) into organic carbon, which they exploit to build cells and provide energy; happily for us animals, free oxygen is their waste product. But plants are finicky, and they preferentially choose CO_2 containing ^{12}C. Thus, when plant life—whether in the form of photosynthesizing microbes, floating algae or tall trees—is abundant, a higher proportion of CO_2 remaining in the atmosphere contains ^{13}C, and atmospheric ^{12}C is measurably lower.

By examining the isotope ratios in samples from before, during and after a mass extinction, investigators can obtain a reliable indicator of the amount of plant life both on land and in the sea. When researchers plot such measurements for the K/T event on a graph, a simple pattern emerges. Virtually simultaneously with the emplacement of the so-called impact layer containing mineralogical evidence of debris, the carbon isotopes shift—^{13}C drops dramatically—for a short time, indicating a sudden die-off of plant life and a quick recovery.

Patterns of Destruction

Carbon 13 (^{13}C) isotopes found in geologic strata suggest longer-acting mechanisms behind two of three ancient extinction events. ^{13}C is more abundant in the atmosphere when land and sea plants are thriving. When plant life dies on a massive scale, ^{13}C drops as a proportion of atmospheric carbon. Comparing ancient samples with a common carbon standard reveals multiple large drops in ^{13}C leading up to the end Permian (*top*) and end Triassic (*middle*) boundaries. The dips imply multiple extinction crises occurring over hundreds of thousands of years. In contrast, a ^{13}C plunge for the period around the Cretaceous-Tertiary boundary (*bottom*) depicts one abrupt ecological cataclysm.



JANET CHAO

This finding is entirely consistent with the fossil record of both larger land plants and the sea's microscopic plankton, which underwent staggering losses in the K/T event but bounced back rapidly.

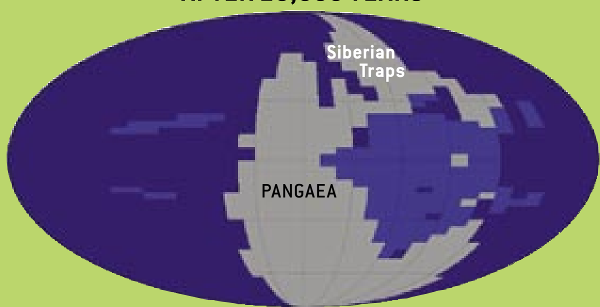
In contrast, the carbon records revealed by my group in early 2005 for the Permian, and more recently for the Triassic, document a very different fate for plants and plankton during those two mass extinctions. In both cases, multiple isotope shifts over intervals exceeding 50,000 to 100,000 years indicate that plant communities were struck down, then re-formed, only to be perturbed again by a series of extinction events [*see box on preceding page*]. To produce such a pattern would take a succession of asteroid strikes, thousands of years apart. But no mineralogical evidence exists for a string of impacts during either time span.

Indeed, further investigation of the evidence has called into question the likelihood of *any* impacts at those two times. No other research groups have replicated the original finding of buckyballs containing extraterrestrial gas at the end Permian boundary. A discovery of shocked quartz from that period has also been recanted, and geologists cannot agree whether purported impact craters from the event in the deep ocean near Australia and under ice in Antarctica are actually craters or just natural rock formations. For the end Triassic, the iridium found is in such low concentrations that it might reflect a small asteroid impact, but nothing of the planet-killing scale seen at the K/T boundary. If impacts are not supported as the cause of these mass extinctions, however, then what did trigger the great die-offs? A new type of evidence reveals that the earth itself can, and probably did, exterminate its own inhabitants.

Slow Poisoning

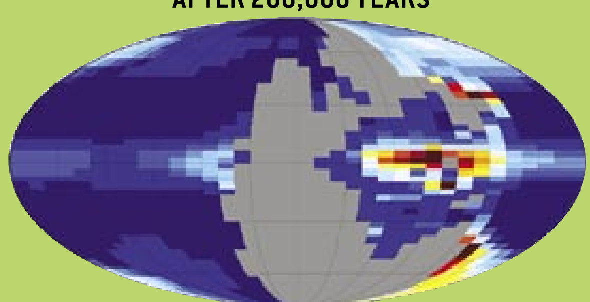
Computer simulations project rising concentrations of toxic hydrogen sulfide and gradual oxygen depletion in the surface waters of the world's oceans at the end of the Permian period. The model by Katja M. Meyer and Lee R. Kump of Pennsylvania State University illustrates how global warming caused by extensive volcanic activity beginning around 251 million years ago in the Siberian Traps region of the Pangaea supercontinent affected the oceans, triggering ecosystem catastrophe.

AFTER 20,000 YEARS

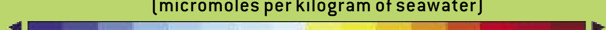


AFTER 80,000 YEARS

AFTER 200,000 YEARS



DISSOLVED HYDROGEN SULFIDE
(micromoles per kilogram of seawater)



KATJA M. MEYER AND LEE R. KUMP *Pennsylvania State University/NASA Astrobiology Institute*; DAVID FIERSTEIN (*opposite page*)

Ghastly Greenhouse

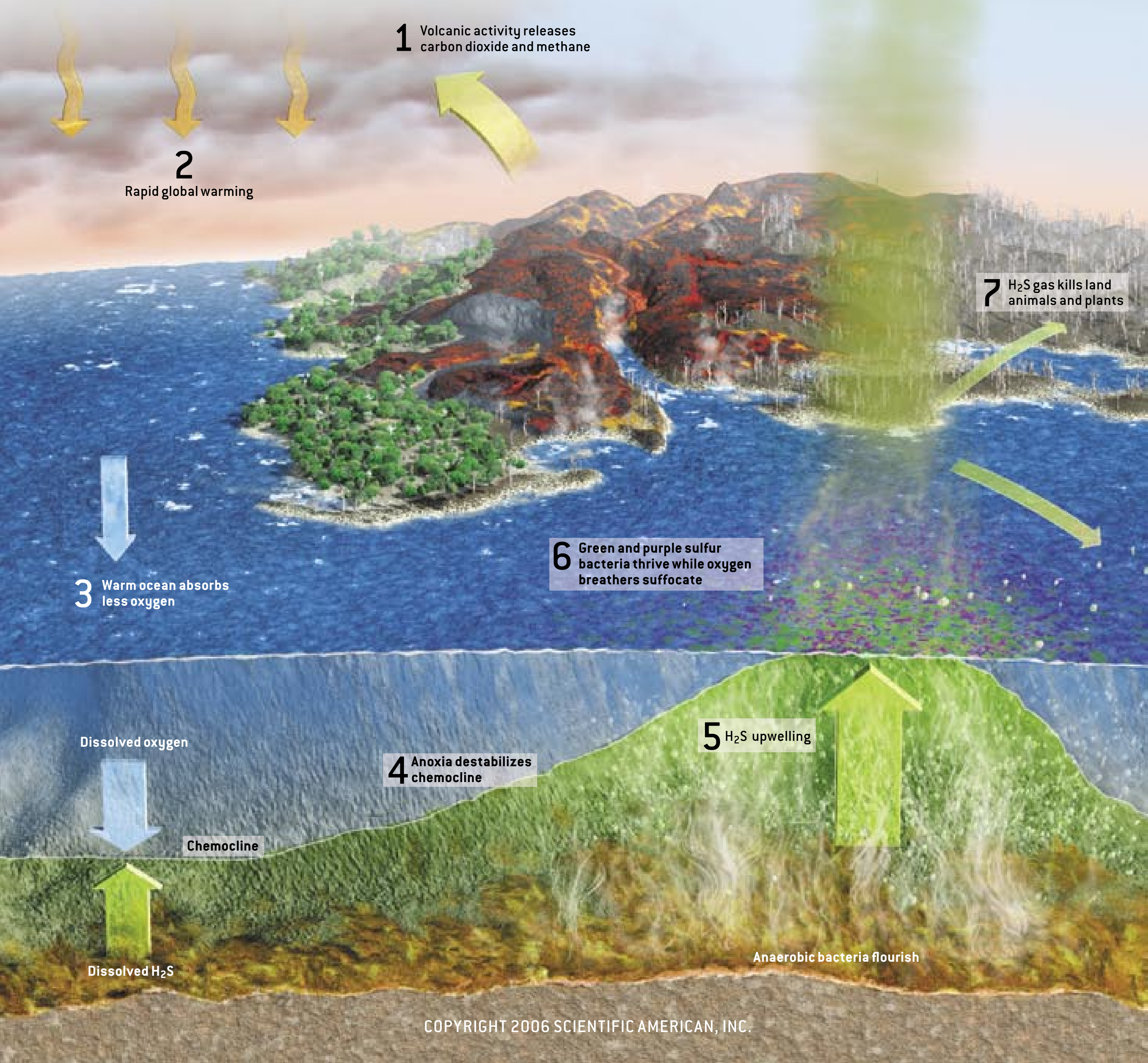
ABOUT HALF A DECADE AGO small groups of geologists began to team up with organic chemists to study environmental conditions at critical times in the earth's history. Their work involved extracting organic residues from ancient strata in search of chemical "fossils" known as biomarkers. Some organisms leave behind tough organic molecules that survive the decay of their bodies and become entombed in sedimentary rocks. These biomarkers can serve as evidence of long-dead life-forms that usually do not leave any skeletal fossils. Various kinds of microbes, for example, leave behind traces of the distinctive lipids present in their cell membranes—traces that show up in new forms of mass spectrometry, a technique that sorts molecules by mass.

This biomarker research was first conducted on rocks predating the history of animals and plants, in part to determine when and under what conditions life first emerged on the earth. But within the past few years scientists began sampling the mass extinction boundaries. And to the great surprise of those doing this work, data from the periods of mass extinction, other than the K/T event, suggested that the world's oceans have more than once reverted to the extremely low oxygen conditions, known as anoxia, that were common before plants and animals became abundant.

Among the biomarkers uncovered were the remains of large numbers of tiny photosynthetic green sulfur bacteria. Today these microbes are found, along with their cousins, photosynthetic purple sulfur bacteria, living in anoxic marine environments such as the depths of stagnant lakes and the Black Sea, and they are pretty noxious characters. For energy, they

KILLER GREENHOUSE EFFECT

A new model for mass extinctions at the end of the Permian period 251 million years ago and the end Triassic 50 million years later explains how intense global warming could trigger deaths in the sea and on land. Trouble begins with widespread volcanic activity that releases enormous volumes of carbon dioxide and methane (*1*). The gases cause rapid global warming (*2*). A warmer ocean absorbs less oxygen from the atmosphere (*3*). Low oxygen (anoxia) destabilizes the chemocline, where oxygenated water meets water permeated with hydrogen sulfide (H_2S) generated by bottom-dwelling anaerobic bacteria (*4*). As H_2S concentrations build and oxygen falls, the chemocline rises abruptly to the ocean surface (*5*). Green and purple photosynthesizing sulfur bacteria, which consume H_2S and normally live at chemocline depth, now inhabit the H_2S-rich surface waters while oxygen-breathing ocean life suffocates (*6*). H_2S also diffuses into the air, killing animals and plants on land (*7*) and rising to the troposphere to attack the planet's ozone layer (*8*). Without the ozone shield, the sun's ultraviolet (UV) radiation kills remaining life (*9*).



HYDROGEN SULFIDE ERUPTIONS off the coast of Namibia appear in this satellite photograph as pale green swirls on the ocean surface. These regular local events, which result from buildup of hydrogen sulfide in sea-bottom sediments, offer a small modern taste of conditions during the global upwellings proposed for several ancient mass extinction periods: a sulfurous smell fills the air, dead fish litter the water, and oxygen-starved lobsters flee onto beaches trying to escape the poison.

oxidize hydrogen sulfide (H_2S) gas, a poison to most other forms of life, and convert it into sulfur. Thus, their abundance at the extinction boundaries opened the way for a new interpretation of the cause of mass extinctions.

Scientists have long known that oxygen levels were lower than today around periods of mass extinction, although the reason was never adequately identified. Large-scale volcanic activity, also associated with most of the mass extinctions, could have raised CO_2 levels in the atmosphere, reducing oxygen and leading to intense global warming—long an alternative theory to the impacts; however, the changes wrought by volcanism could not necessarily explain the massive marine extinctions of the end Permian. Nor could volcanoes account for plant deaths on land, because vegetation would thrive on increased CO_2 and could probably survive the warming.

But the biomarkers in the oceanic sediments from the latest part of the Permian, and from the latest Triassic rocks as well, yielded chemical evidence of an ocean-wide bloom of the H_2S-consuming bacteria. Because these microbes can live only in an oxygen-free environment but need sunlight for their photosynthesis, their presence in strata representing shallow marine settings is itself a marker indicating that even the surface of the oceans at the end of the Permian was without oxygen but was enriched in H_2S.

In today's oceans, oxygen is present in essentially equal concentrations from top to bottom because it dissolves from the atmosphere into the water and is carried downward by ocean circulation. Only under unusual circumstances, such as those that exist in the Black Sea, do anoxic conditions below the surface permit a wide variety of oxygen-hating organisms to thrive in the water column. Those deep-dwelling anaerobic microbes churn out copious amounts of hydrogen sulfide, which also dissolves into the seawater. As its concentration builds, the H_2S diffuses upward, where it encounters oxygen diffusing downward. So long as their balance remains undisturbed, the oxygenated and hydrogen sulfide–saturated waters stay separated, and their interface, known as the chemocline, is stable. Typically the green and purple sulfur bacteria live in that chemocline, enjoying the supply of H_2S from below and sunlight from above.

Yet calculations by geoscientists Lee R. Kump and Michael A. Arthur of Pennsylvania State University have shown that if oxygen levels drop in the oceans, conditions begin to favor the deep-sea anaerobic bacteria, which proliferate and produce greater amounts of hydrogen sulfide. In their models, if the deepwater H_2S concentrations were to increase beyond a critical threshold during such an interval of oceanic anoxia, then the chemocline separating the H_2S-rich deepwater from oxygenated surface water could have floated up to the top abruptly. The horrific result would be great bubbles of toxic H_2S gas erupting into the atmosphere.

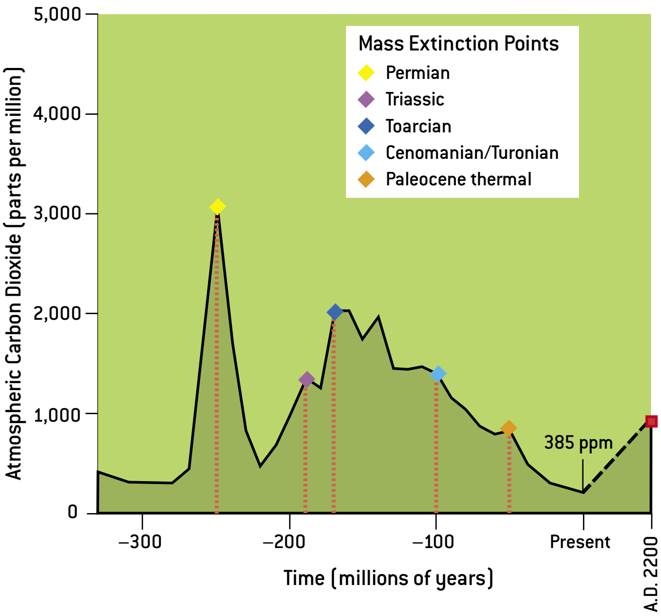
Their studies indicate that enough H_2S was produced by such ocean upwellings at the end of the Permian to cause extinctions both on land and in the sea [*see box on page 68*]. And this strangling gas would not have been the only killer. Models by Alexander Pavlov of the University of Arizona show that the H_2S would also have attacked the planet's ozone shield, an atmospheric layer that protects life from the sun's ultraviolet (UV) radiation. Evidence that such a disruption of the ozone layer did happen at the end of the Permian exists in fossil spores from Greenland, which display deformities known to result from extended exposure to high UV lev-

THE AUTHOR

PETER D. WARD is a professor in the University of Washington's biology department and its earth and space sciences division, where he investigates both realms. His terrestrial research centers on ancient mass extinction events as well as the evolution and ultimate extinction of the nautiluslike marine animals known as ammonites, which he described in his first article for *Scientific American* in October 1983. Ward also applies principles gleaned from studying the earth's earliest life-forms to research for the NASA Astrobiology Institute into potential habitats for life elsewhere. He discussed those environments in an October 2001 *Scientific American* article, "Refuges for Life in a Hostile Universe," written with Guillermo Gonzalez and Donald Brownlee, as well as in a popular book co-authored with Brownlee, *Rare Earth: Why Complex Life Is So Uncommon in the Universe* (Springer, 2000).

JACQUES DESCLOITRES *MODIS Rapid Response Team, NASA Goddard Space Flight Center*

HEADED FOR ANOTHER EXTINCTION?



ATMOSPHERIC CARBON DIOXIDE (CO_2) was high during ancient mass extinctions, supporting a role for global warming in those events. Today CO_2 stands at 385 parts per million (ppm) and is projected to climb by 2 to 3 ppm annually. If this trend continues, by the end of the next century atmospheric CO_2 would approach 900 ppm—just below levels during the Paleocene thermal extinction 54 million years ago.

els. Today we can also see that underneath "holes" in the ozone shield, especially in the Antarctic, the biomass of phytoplankton rapidly decreases. And if the base of the food chain is destroyed, it is not long until the organisms higher up are in desperate straits as well.

Kump and Arthur estimate that the amount of H_2S gas entering the late Permian atmosphere from the oceans was more than 2,000 times the small amount given off by volcanoes today. Enough of the toxic gas would have permeated the atmosphere to have killed both plants and animals—particularly because the lethality of H_2S increases with temperature. And several large and small mass extinctions seem to have occurred during short intervals of global warming. That is where the ancient volcanic activity may have come in.

Around the time of multiple mass extinctions, major volcanic events are known to have extruded thousands of square kilometers of lava onto the land or the seafloor. A by-product of this tremendous volcanic outpouring would have been enormous volumes of carbon dioxide and methane entering the atmosphere, which would have caused rapid global warming. During the latest Permian and Triassic as well as in the early Jurassic, middle Cretaceous and late Paleocene, among other periods, the carbon-isotope record confirms that CO_2 concentrations skyrocketed immediately before the start of the extinctions and then stayed high for hundreds of thousands to a few million years.

But the most critical factor seems to have been the oceans. Heating makes it harder for water to absorb oxygen from the atmosphere; thus, if ancient volcanism raised CO_2 and lowered the amount of oxygen in the atmosphere, and global warming made it more difficult for the remaining oxygen to penetrate the oceans, conditions would have become amenable for the deep-sea anaerobic bacteria to generate massive upwellings of H_2S. Oxygen-breathing ocean life would have been hit first and hardest, whereas the photosynthetic green and purple H_2S-consuming bacteria would have been able to thrive at the surface of the anoxic ocean. As the H_2S gas choked creatures on land and eroded the planet's protective shield, virtually no form of life on the earth was safe.

Kump's hypothesis of planetary killing provides a link between marine and terrestrial extinctions at the end of the Permian and explains how volcanism and increased CO_2 could have triggered both. It also resolves strange findings of sulfur at all end Permian sites. A poisoned ocean and atmosphere would account for the very slow recovery of life after that mass extinction as well.

Finally, this proposed sequence of events pertains not only to the end of the Permian. A minor extinction at the end of the Paleocene epoch 54 million years ago was already—presciently—attributed to an interval of oceanic anoxia somehow triggered by short-term global warming. Biomarkers and geologic evidence of anoxic oceans suggest that is also what may have occurred at the end Triassic, middle Cretaceous and late Devonian, making such extreme greenhouse-effect extinctions possibly a recurring phenomenon in the earth's history.

Most troubling, however, is the question of whether our species has anything to fear from this mechanism in the future: If it happened before, could it happen again? Although estimates of the rates at which carbon dioxide entered the atmosphere during each of the ancient extinctions are still uncertain, the ultimate levels at which the mass deaths took place are known. The so-called thermal extinction at the end of the Paleocene began when atmospheric CO_2 was just under 1,000 parts per million (ppm). At the end of the Triassic, CO_2 was just above 1,000 ppm. Today with CO_2 around 385 ppm, it seems we are still safe. But with atmospheric carbon climbing at an annual rate of 2 ppm and expected to accelerate to 3 ppm, levels could approach 900 ppm by the end of the next century, and conditions that bring about the beginnings of ocean anoxia may be in place. How soon after that could there be a new greenhouse extinction? That is something our society should never find out. SA

MORE TO EXPLORE

Rivers in Time: The Search for Clues to Earth's Mass Extinctions. Peter D. Ward. Columbia University Press, 2002.

Abrupt and Gradual Extinction among Late Permian Land Vertebrates in the Karoo Basin, South Africa. Peter D. Ward et al. in *Science,* Vol. 307, pages 709–714; February 4, 2005.

Photic Zone Euxinia during the Permian-Triassic Superanoxic Event. Kliti Grice et al. in *Science,* Vol. 307, pages 706–709; February 4, 2005.

Massive Release of Hydrogen Sulfide to the Surface Ocean and Atmosphere during Intervals of Oceanic Anoxia. Lee R. Kump, Alexander Pavlov and Michael A. Arthur in *Geology,* Vol. 33, No. 5, pages 397–400; May 2005.

JANET CHAO

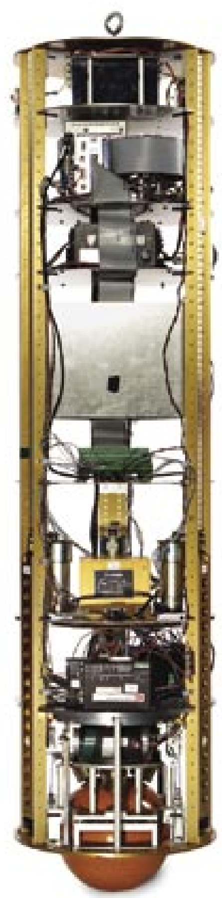
A new mode of locomotion will enable mobile robots to stand tall and move gracefully through busy everyday environments

BALLBOTS

By Ralph Hollis

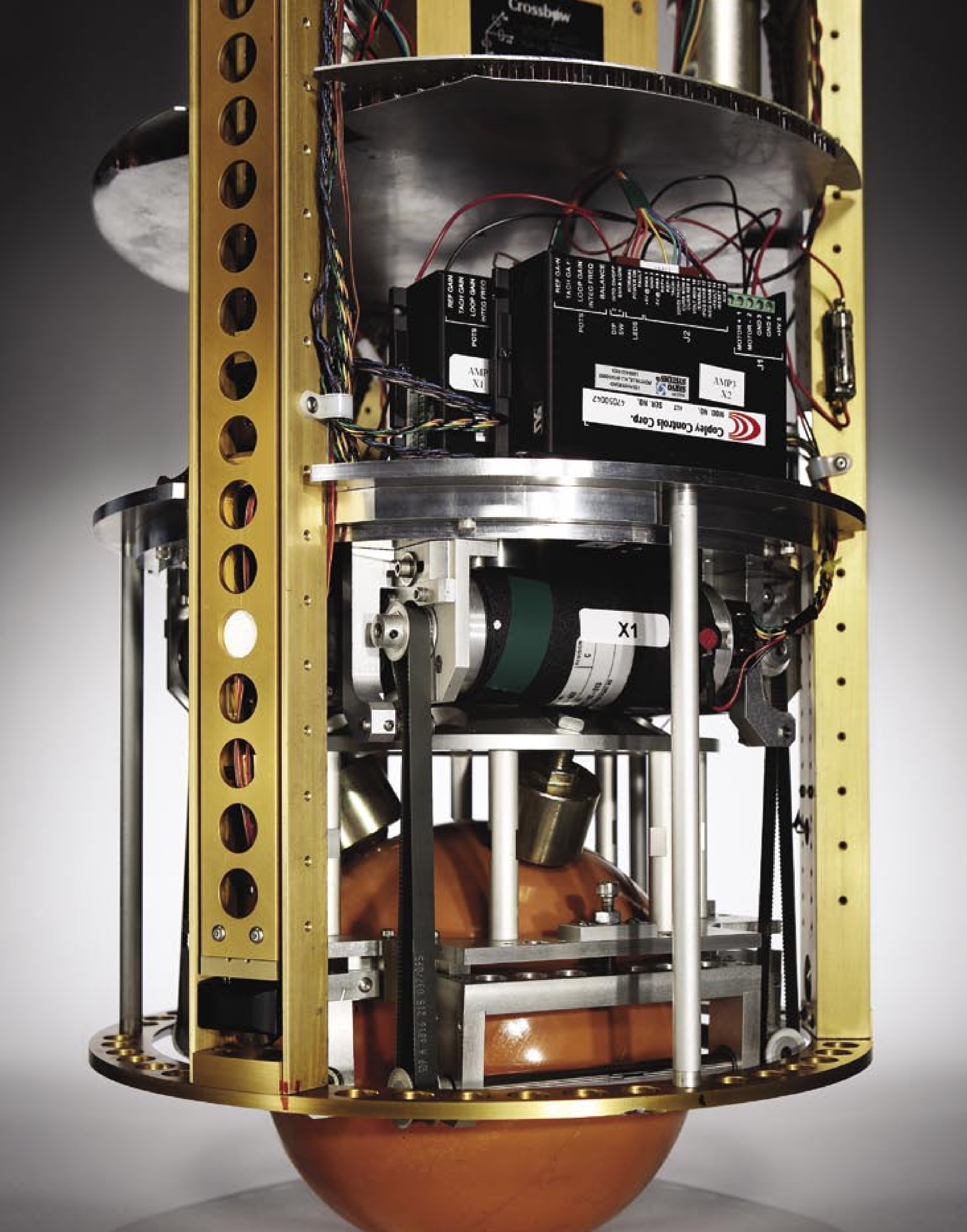
The dream of intelligent, mobile robots that assist people during their day-to-day activities in homes, offices and nursing facilities is a compelling one. Although a favorite subject of science-fiction writers and robotics researchers, the goal seems always to lie well off in the future, however. Engineers have yet to solve fundamental problems involving robotic perception and world modeling, automated reasoning, manipulation of objects and locomotion.

Researchers have produced robots that, while falling far short of the ideal, can do some remarkable things. In 2002 one group dropped off a robot at the entrance to the annual meeting of the American Association for Artificial Intelligence in Edmonton, Alberta. The clever machine soon found its way to the registration booth, signed up for the conference, was assigned a lecture room, proceeded to that location and finally presented a brief talk about itself at the appointed hour. Some robots have in the meantime served effectively as interactive museum tour guides, whereas others show promise as nursing home assistants. Computer scientists and engineers have also equipped mobile systems with arms and hands for manipulating objects. All these experimental devices travel



MOBILE ROBOTICS takes a different path with the ballbot's unique single, spherical drive wheel design.

BRIAN MARANAN PINEDA


Crossbow
Copley Controls Corp.
AMP3
X2
X1

about on bases supported by three or four wheels. Designers call this configuration "statically stable" because it keeps the robots upright even at rest.

Robots tall enough to interact effectively in human environments have a high center of gravity and must accelerate and decelerate slowly, as well as avoid steep ramps, to keep from falling over. To counter this problem, statically stable robots tend to have broad bodies on wide wheelbases, which greatly restricts their mobility through doorways and around furniture or people.

Several years ago I decided to sidestep the need for large wheelbases by designing and building a tall, skinny and agile robot that balances on, and is propelled by, a single spherical wheel. Such a simple machine, with its high center of gravity, would be able to move quickly in any direction. The system would rely on active balancing and thus be "dynamically stable"—that is, it would remain erect only if it made continual corrections to its body attitude. I realized this design would constitute a hitherto unstudied class of wheeled mobile robots. For lack of anything better I called it a ballbot.

My students and I have operated our ballbot now for more than a year, studying its stability properties and suitability for operating in human environments. During that time, many visitors to our laboratory have found its uncanny ability to balance and roam about on a single spherical wheel to be quite remarkable.

Maintaining Balance

WE HUMANS KEEP BALANCE with help from the vestibular senses in our inner ears. This information is combined with input from other senses, such as vision, to control muscles in our legs and feet to enable us to stand upright without falling down. A ballbot maintains equilibrium in a somewhat analogous fashion. First, the machine must have some goal to achieve, such as to remain in one place or to move in a straight line between two locations. Second, it must always know the direction of gravity's pull and be able to measure the orientation of its body with respect to this vertical reference. Third, it must have means to rotate the ball in any direction and to measure its travel along the floor. Finally, the ballbot must have a method, or control policy, that processes the sensor data it measures to generate commands for ball rotation that attempt to satisfy the goals.

> **Many visitors find its uncanny ability to balance and roam about on a single spherical wheel to be quite remarkable.**

Solving the "problem of the vertical" has proved to be a challenging exercise throughout history [*see box on page 76*]. Our solution takes advantage of tremendous recent advances in computing, fiber optics and microelectromechanical systems (MEMS) that have enabled the production of low-cost devices that emulate the function of the traditional spinning gyroscope.

We use a system that features three fiber-optic gyroscopes mounted orthogonally (at right angles to one another) in a box that is rigidly attached to the ballbot body [*see box on opposite page*]. These gyroscopes contain no rotating masses. Each gyroscope features a light source, a detector and a coil of optical fiber. Light waves travel around the coil in opposite directions and interfere with one another at the detector. During operation, the ballbot body, with its three gyroscopic, angular-motion sensors, rotates in various directions, but the light waves inside them travel at a fixed speed regardless of any movement. Accordingly, a small path difference between the clockwise- and counterclockwise-propagating waves results in each sensor. In each case, the path difference causes the interference fringes at the detector to shift, producing an output that is proportional to angular velocity, an effect noted by French physicist Georges Sagnac as far back as 1913. A small computer integrates the three angular velocities to produce pitch (forward/backward tilt), roll (left/right tilt) and yaw (rotation around the vertical) angles taken by the robot's body.

To report the correct vertical orientation, all gyroscopes must take into account the earth's rotation. They are also subject to numerous other small effects that cause errors and drift over time. Our system incorporates three MEMS accelerometers, set orthogonally in the same box alongside the gyroscopes. As the ballbot moves around, these sensors report the resulting instantaneous acceleration values for each orientation, which the computer then combines to yield an overall acceleration direction and magnitude that can be averaged over time. (The accelerometers' readings cannot be used directly for balancing.) The outcome is a reliable long-term indicator of the direction of gravity that the system uses to correct the drift of the fiber-optic gyroscopes.

Moving with the Ball

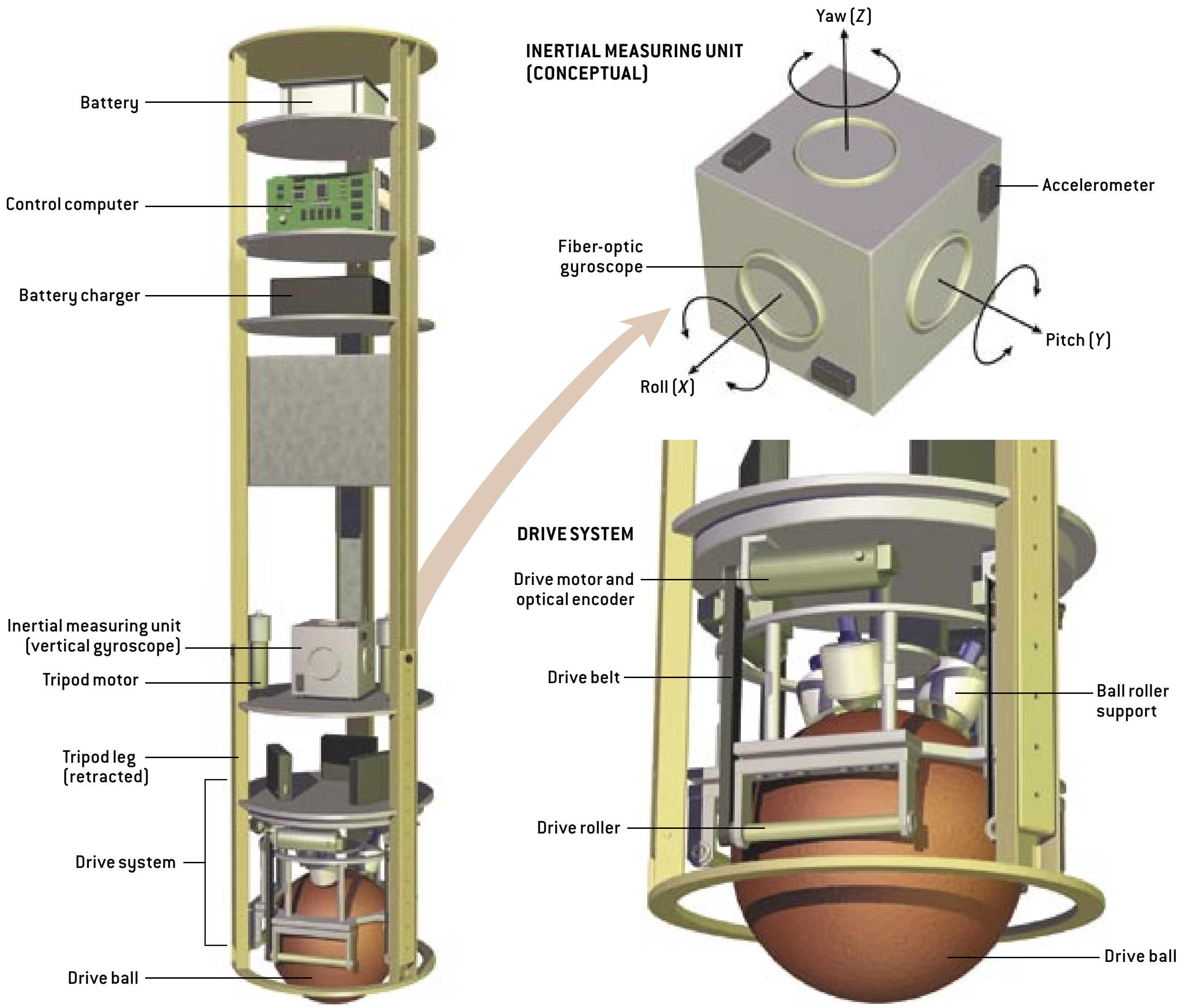
SEVERAL METHODS EXIST for driving a ball in various directions using motors. We strove for simplicity in our design for the ballbot's drive mechanism. When one moves a mechanical comput-

Overview/*Mobile Robots*

- To interact with people in their everyday environments, intelligent mobile robots will need to stand tall, as well as to move surely and gracefully.
- Most current experimental mobile robots feature wide wheelbases, which hinder their movements through cramped, chaotic human settings.
- A ballbot—a tall, thin robot that travels about on a ball-shaped wheel that enables it to move rapidly in any direction—may provide the flexible locomotive capabilities that future robots will need to aid people in their daily lives.

BALLBOT ARCHITECTURE

In some ways, a ballbot (*left*) resembles a ballpoint pen that is five feet tall. The fiber-optic gyroscopes and accelerometers (*top right*), which are mounted at right angles to one another to sense motion in the pitch, roll and yaw directions, generate the vertical orientation data the computer control policy needs to determine how to maintain balance [*see box on next page for explanation of underlying principles*]. The drive ball mechanism (*bottom right*), which operates something like an inverse computer mouse, provides the ballbot's motive force. Motorized drive rollers turn the ball, and optical encoders measure the ballbot's travel. To stay upright when shut down, the machine deploys its tripod legs.



er mouse about on the desktop, the rubber-coated ball on the underside causes a pair of orthogonally mounted rollers to turn. The measured rotation of the rollers provides input to the computer to traverse the cursor across the screen. Just the opposite happens in the ballbot: output from the ballbot's computer commands a set of motors to turn rollers that rotate the ball, thus causing the robot to travel in any direction along the floor. It is essentially an "inverse mouse ball" drive. Currently motors actuate the ball in the pitch and roll directions. An additional motor (not yet installed) will rotate the body in yaw, which will allow the ballbot to face in any direction.

Much as a circus clown might perch atop a ball, the ballbot's body stands atop the ball wheel. The ball is a hollow aluminum sphere covered with a thick layer of polyurethane rubber. Such a drive scheme exhibits frictional and damping behavior because sliding always occurs between the ball and rollers, for which compensation must be made. Three ball bearings between the ball and body support the body's weight.

To infer ball rotation and hence travel distance, we used optical encoders that are fitted to each of the drive motors. Each encoder has a fixed light source opposite a light detector. A transparent, rotating mask (with many fine opaque stripes) attached to the motor shaft sits between them. As the motor turns, the mask rotates, causing the striped pattern to alternately block and transmit the light beam. The ballbot's

GEORGE RETSECK

The Problem of the Vertical

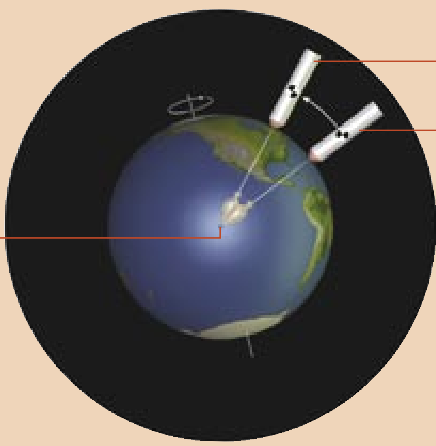
Finding the up/down orientation, what early aviators called the problem of the vertical, continues to be difficult even today. A plumb bob hanging from a string reveals the vertical, but a ballbot equipped with such a pendulum reference would become confused because motion (say, from position A to B, below) would cause the bob to swing to and fro.

Alternatively, the ballbot could rely on a gyroscope. The gyro's wheel would be supported by gimbals, which would allow its axis to point arbitrarily. By driving the wheel with a motor, it could be spun rapidly with its axis aligned vertically before the ballbot began to operate. The inertia of such a gyro would keep it pointing in the same direction regardless of movement. Equipping the gimbals with angle sensors would allow measurement of the body's forward/backward (pitch) and its left/right (roll) attitudes. This approach has problems, however. The gyro's axis would remain fixed in space while the earth rotates and hence would depart from the vertical.

German engineer Maximilian Schuler first formulated a solution to this problem in 1923 by imagining a pendulum string long enough to reach the center of the earth. Such a long string would always point downward regardless of motion. This pendulum would, in fact, have a period of about 84.4 minutes, the so-called Schuler period, which corresponds to the earth's orbital period at its surface on the equator. He showed how small torques exerted on a gyroscope could increase the period of a short, practical pendulum to 84.4 minutes (and thus make it behave like a Schuler pendulum), which would keep it oriented along the direction of gravity.

The ballbot could, in theory, use such a gyro with a short pendulum. As the ballbot moves, the directions of the pendulum's swing could be measured over time and averaged to yield a value that faithfully represented the vertical (because the lateral accelerations would cancel out over time, leaving gravity dominant). The result could be used to exert torques on the gyro to make it stay vertical.

We opted for another solution. Our ballbot uses fiber-optic gyroscopes and microelectromechanical accelerometers that together emulate the functions of a mechanical gyro and pendulum that behaves like a Schuler pendulum. The result is a gravity-seeking, or "vertical," gyro that serves as a reference for balancing. *—R.H.*



GEORGE RETSECK

THE AUTHOR

RALPH HOLLIS is a research professor at Carnegie Mellon University's Robotics Institute, with an appointment in the department of electrical and computer engineering. He received his Ph.D. in solid-state physics from the University of Colorado at Boulder in 1975. Before joining Carnegie Mellon in 1993, Hollis worked at North American Aviation and the IBM Thomas J. Watson Research Center. His current research focuses on agent-based microassembly of electromechanical products, human-computer interaction through the sense of touch, and dynamically stable mobile robots.

main computer counts these events to measure ball rotation and thus distance traveled.

Ball Control

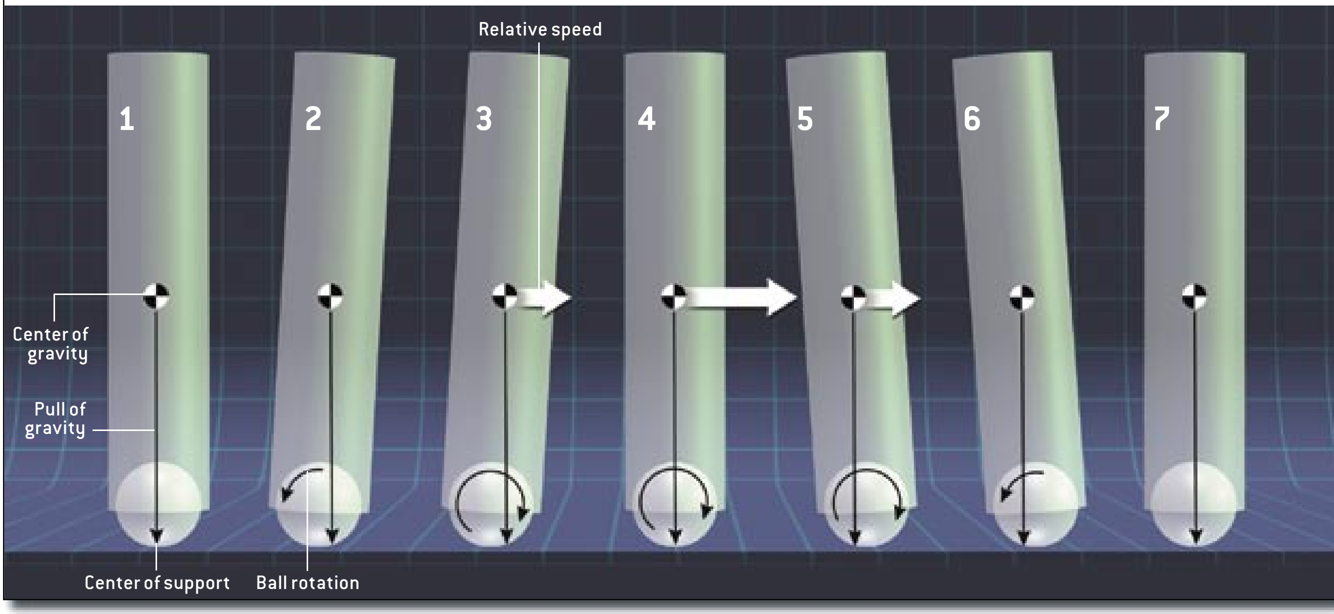
SIMPLY STATED, the ballbot uses its knowledge of the vertical to determine how to rotate its ball to balance and move about. Fortunately, the ballbot is fundamentally an inverted pendulum, a mechanism that physicists have studied extensively. We use the techniques of optimal control theory to find a strategy or policy for driving the ballbot to its goal while simultaneously minimizing the effort it takes to get there. The ballbot has eight internal states that the policy must take into account: four for its forward/backward motion and four for its left/right motion. For each of these directions, the system measures or infers (from the onboard sensors) the robot's position and speed, and the tilt and tilt rate of the body.

We employ a simplified linear mathematical model to describe the ballbot's dynamics. Rudolf Kalman, a Hungarian-American mathematical system theorist, invented in 1960 an elegant method for deriving control policies for such systems, which he called the linear quadratic regulator. This approach considers the measurements of the system's internal states to be proportional to the values of the states themselves. Further, it assumes that the states change over time at a rate proportional to the values of the states plus a proportional contribution of any control actions that might occur, such as motor torques. Kalman's technique cleverly minimizes an integral function over time that includes a quadratic measure of the states plus a quadratic measure of the control actions. Its solution yields a final set of constants, which, when multiplied by each of the internal states, gives a recommended, or optimal, control action for the ballbot to take at each moment in time. These calculations run several hundred times a second in the ballbot's main computer.

When the ballbot's goal is to stand still, its control policy tries to simultaneously drive the body's position and speed as well as its tilt and tilt rate to

TRAVELING FROM HERE TO THERE

To maintain balance when still, the ballbot must keep its center of gravity directly over its center of support (*1*). Orientation sensors determine the vertical direction, which the machine then compares with its current attitude. During movement, the ballbot manipulates its center of gravity to best effect. To go from one point to another on level ground, for example, the drive ball first rotates slightly in the direction opposite to the intended direction of travel (*2*), which tilts the body forward a bit to initiate the move. Next, the ball spins in the direction of motion to accelerate ahead (*3*). While the ballbot is at constant velocity, the body must remain nearly vertical (*4*). The opposite actions must occur to decelerate the machine (*5*) and then prepare it to halt (*6*), which together bring it to a stop (*7*). When traversing inclines, the body must lean into slopes to keep its equilibrium.



zero in each direction, while minimizing the actions needed to do so. When its objective is to go from one place to another, the control policy automatically institutes a retrograde ball rotation to establish a body tilt, allowing it to accelerate forward. As the goal position is approached, the ball automatically speeds up to reverse the tilt and bring the ballbot to rest [*see box above*].

Moving Ahead

WE HAVE JUST BEGUN to experiment with the ballbot, interacting with it over a wireless radio link. We plan to add a pair of arms as well as a head that pans and tilts, with a binocular vision system and many other sensors in an effort to develop the machine into a capable robot with a significant degree of autonomy. Our goals are to understand how well such robots can perform around people in everyday settings and to compare quantitatively its performance, safety and navigation abilities with those of traditional, statically stable robots. Our hypothesis is that the latter may turn out to be an evolutionary dead end when it comes to operating in such environments.

We are not alone in betting on the notion of dynamically stable robots. Other research groups have produced two-wheeled robots that are dynamically stable in the pitch direction but statically stable in the roll orientation. Although these robots are not omnidirectional like a ballbot is, they show promise for agile mobility—especially outdoors.

It may turn out that dynamically stable biped robots, perhaps in humanoid form, will have the long-term edge—particularly for their ability to deal with stairways. Research teams worldwide are working intensively to develop these complex and often expensive machines. Meanwhile it would seem that ballbots will serve as interesting and effective platforms for studying how mobile robots can interact dynamically and gracefully with humans in the places where people live. SA

MORE TO EXPLORE

For a discussion of gyroscopic principles, see **Inventing Accuracy: A Historical Sociology of Nuclear Missile Guidance.** Reprint edition. Donald MacKenzie. MIT Press, 1993.

A Dynamically Stable Single-Wheeled Mobile Robot with Inverse Mouse-Ball Drive. T. B. Lauwers, G. A. Kantor and R. L. Hollis in *Proceedings of the 2006 IEEE International Conference on Robotics and Automation (ICRA '06),* May 2006.

One Is Enough! Tom Lauwers, George Kantor and Ralph Hollis in *Robotics Research: The Twelfth International Symposium.* Springer Tracts in Advanced Robotics (in press).

Ballbot information (including demonstration videos): **www.msl.ri.cmu.edu/projects/ballbot/**

GRACE: The Social Robot: **www.palantir.swarthmore.edu/GRACE/**

Information on fiber-optic gyros: **http://leoss.feri.uni-mb.si/dip_vedran.html**

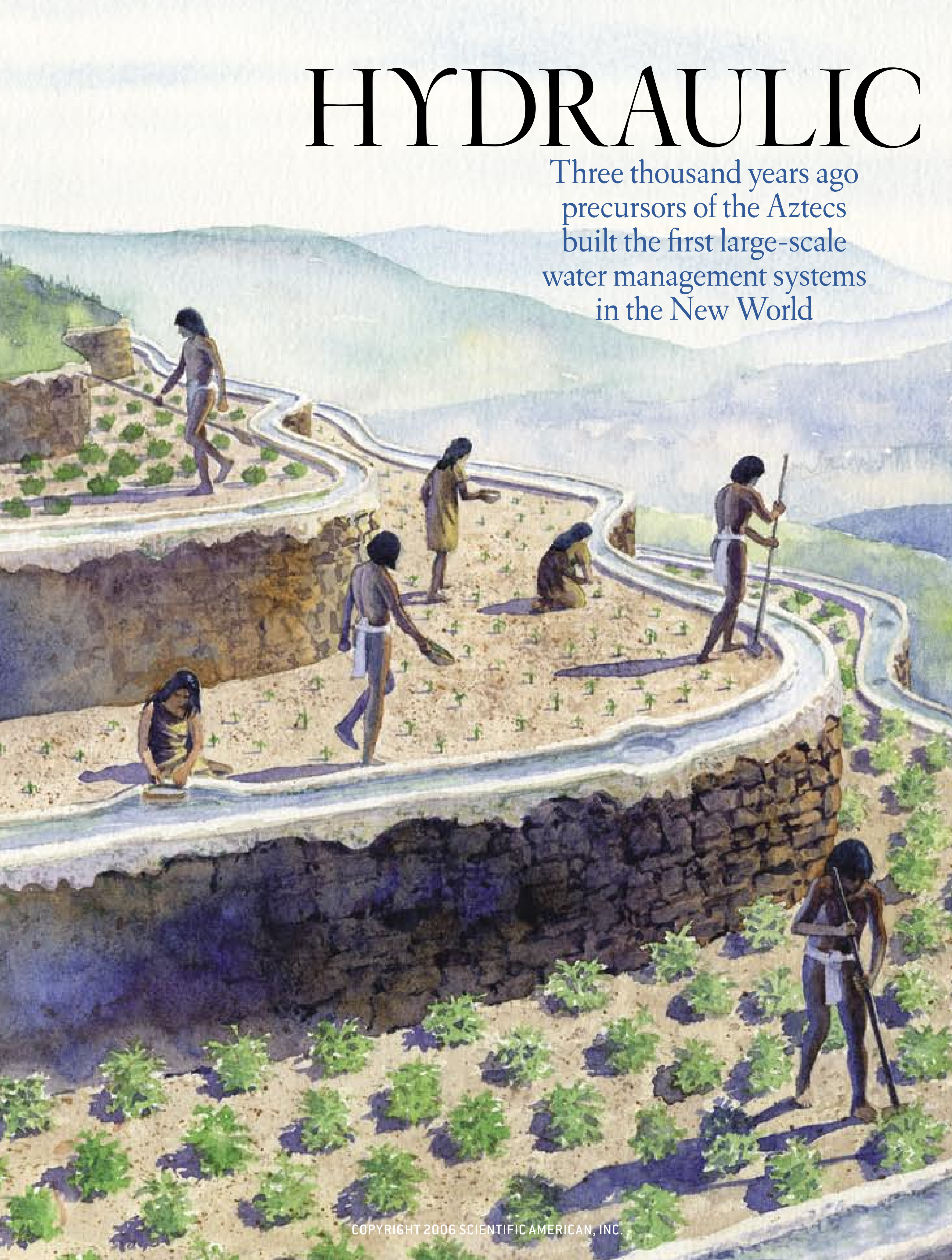
Information on the linear quadratic regulator: **http://en.wikipedia.org/wiki/Linear-quadratic_regulator**

Information on MEMS accelerometers: **www.designnews.com/article/CA294124.html**

GEORGE RETSECK

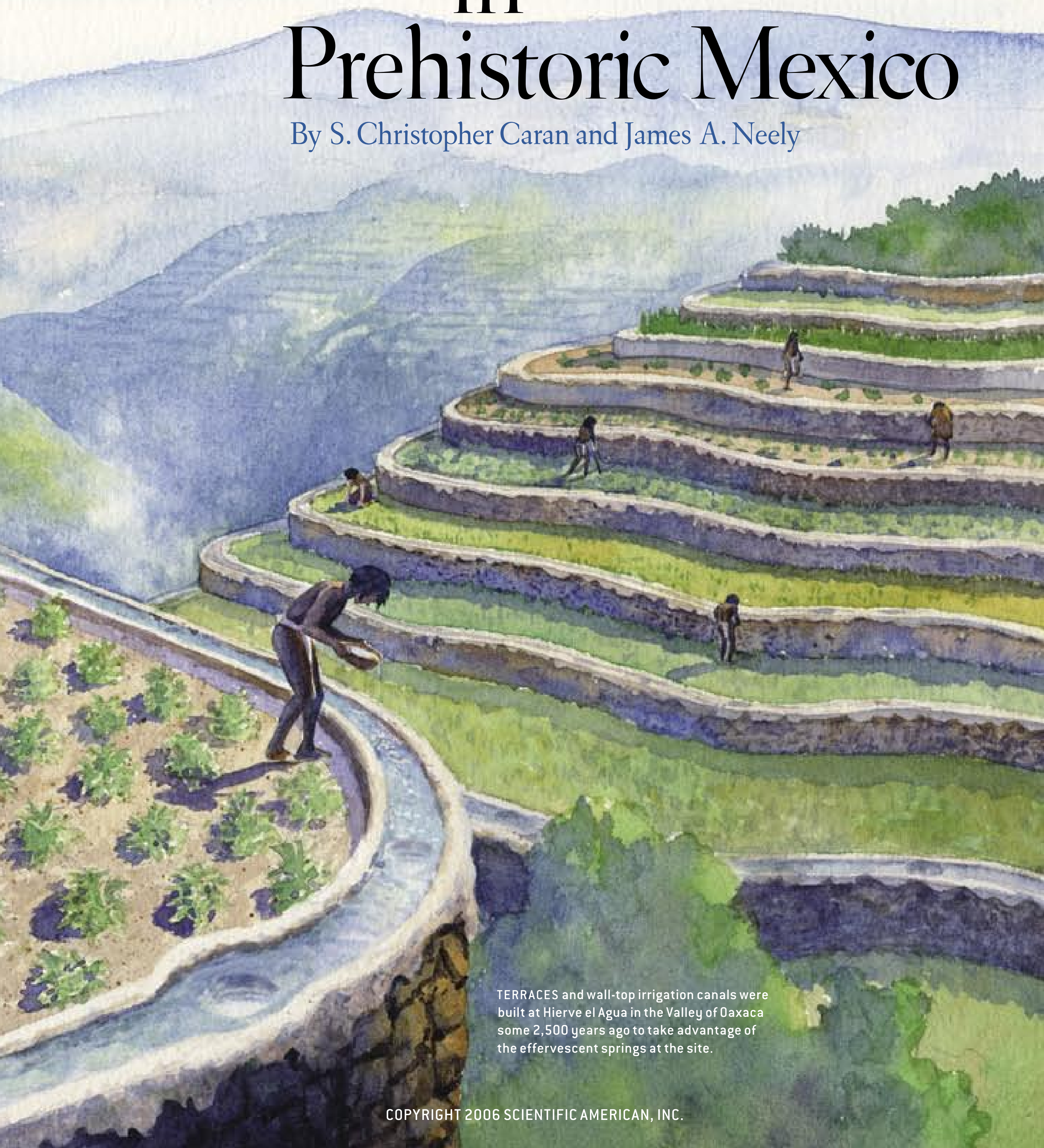
HYDRAULIC

Three thousand years ago precursors of the Aztecs built the first large-scale water management systems in the New World



ENGINEERING in Prehistoric Mexico

By S. Christopher Caran and James A. Neely



TERRACES and wall-top irrigation canals were built at Hierve el Agua in the Valley of Oaxaca some 2,500 years ago to take advantage of the effervescent springs at the site.

The prehistoric farmers of southern Mexico must have longed for a miracle. A tropical climate made their fertile valleys nearly ideal for planting, despite elevations approaching 2,000 meters, and heavy rains ensured bountiful crops during the six-month monsoon season. Under such favorable conditions, this region became the cradle of New World agriculture and the birthplace of corn. Yet these early agriculturalists faced one crucial limitation: during half the year, the weather was too dry for farming. With a year-round water supply, their hand-tilled fields might yield two or even three harvests annually. But how could the farmers get more water?

Their solution was not a miracle but a marvel of human ingenuity: large-scale engineering projects designed to store and transport water. From modest origins that left few traces, construction gradually progressed to a monumental scale. The Purrón Dam, for example, which was built in the Tehuacán Valley starting around 750 B.C., measured 400 meters long, 100 meters wide and nearly 25 meters high. Workers transported by hand, a few kilograms at a time, some 2.64 million cubic meters of earth. This dam probably remained the largest water retention structure in the Americas until the 18th century. Nearby, ancient engineers built thousands of kilometers of canals and aqueducts that predated the arrival of Europeans in Mexico by two millennia. They diverted water from springs and streams, conveying it across drainage divides, around canyons and down steep slopes. Other resourceful inventions collected rainwater from buildings and plazas. The people of southern Mexico exploited virtually every source of water in their environment.

Many of the major water collection and irrigation structures have survived in excellent condition for 1,500 to nearly 3,000 years—testimony to their superb design and construction. These accomplishments are extraordinary by any measure but are particularly notable because the builders had no metal tools, no wheeled transport and no draught animals. Even the oldest surviving features represent a high level of technological innovation—and hint at sophisticated management to maintain this massive infrastructure. Although prehistoric water management systems have been discovered at a number of locations in Mexico, a close look at two of them—the extensive canals of the Tehuacán Valley and the fantastic terraced irrigation network in the Valley of Oaxaca—will illuminate the ingenuity of the ancient hydraulic engineers.

Modern investigation of these two sites began in the late 1960s and early 1970s, when two legendary figures in archaeology led major surveys: Richard S. ("Scotty") MacNeish in the Tehuacán Valley and Kent V. Flannery in the Valley of Oaxaca. One of us (Neely) was then a young graduate student and was privileged to participate in both studies. Water management was not the focus of either survey, and despite the obvious significance of these systems, a long hiatus followed the original modest assessments. Neely's fascination with prehistoric water management persisted, however, and in the late 1980s he invited a geologist (Caran, co-author of this article) to join him in a more thorough analysis of these waterworks. Our findings, as will be seen, were astounding.

WATER MANAGEMENT systems were constructed at various sites throughout southern Mexico during the prehistoric period. The earliest known example, a well at San Marcos Necoxtla, is 10,000 years old. Construction of later systems continued until the arrival of the Europeans.

Overview/*Ancient Irrigation*

- Prehistoric farmers in southern Mexico faced a water shortage six months of the year.
- To overcome this impediment to year-round cultivation, they developed ingenious ways to bring water to their crops. Their inventions included dams, wells, canals, aqueducts and terraced fields.
- Two of these systems illustrate the resourcefulness of these early engineers—a network of canals in the Tehuacán Valley and, in the Valley of Oaxaca, terraced gardens irrigated from canals atop the retaining walls.

HIERVE EL AGUA today still shows clear traces of the ancient terraced fields, here bordering one of the effervescent spring pools.

COURTESY OF JAMES A. NEELY

Canals, Aqueducts and *Tecoatles*

THE CANAL NETWORK in southern Mexico's Tehuacán Valley turned out to be the largest known prehistoric water management system in the New World. The total length of these canals spans more than 1,200 kilometers. They provided water for 330 square kilometers of cropland—an area almost the size of the Gaza Strip—and they did so 2,500 years ago. The irrigators created a canal by excavating a channel in the soil, probably building small levees on either side. Each canal carried water from an uphill source to a lower lying field, often following a meandering course to keep a gentle gradient of two degrees or less.

Most irrigation water was diverted from large springs. The springwaters were rich in dissolved minerals, particularly calcite, which is a form of calcium carbonate. These minerals preserved the canals by forming a leakproof lining but ultimately threatened their long-term survival. As springwater flowed through a canal, evaporation and changes in pressure and temperature caused the chemicals to become so concentrated that a thin layer of minerals crystallized on the canal's interior surface. The mineral contribution of each liter of water was minuscule, but flow through a major canal probably exceeded half a million liters daily. Layer on layer hardened into a stony coating known as calcareous travertine, similar to most stalactites and stalagmites in caves. These layers accumulated at the average rate of one centimeter a year, or one meter a century.

Deposition was so extensive that many of the canals eventually began to fill in. Flow continued, however, because travertine was deposited wherever the water overflowed as well as on the canal's floor. As a result, the walls of the canal became elevated, forming dikes that generally held the water in, even above ground level. Water continued to overflow occasionally, depositing mineral layers far outside the original canal. In this way, a former small channel in the soil became a ridge as much as five meters high and 30 meters wide at its base, with a canal along its crest [*see illustration on next page*]. Perhaps aided by periodic maintenance, a canal could thus retain its U-shaped cross section and continue to rise and function.

The rocky nature and elongate, sinuous form of these "fossilized" canals inspired their name in the language of the Aztecs: *tecoatl,* or "stone snake." Kilometers-long *tecoatles* literally transformed the landscape, creating barricades that affected the alignment of roads and the design of cities and towns from the prehistoric and the Spanish colonial periods to the present day.

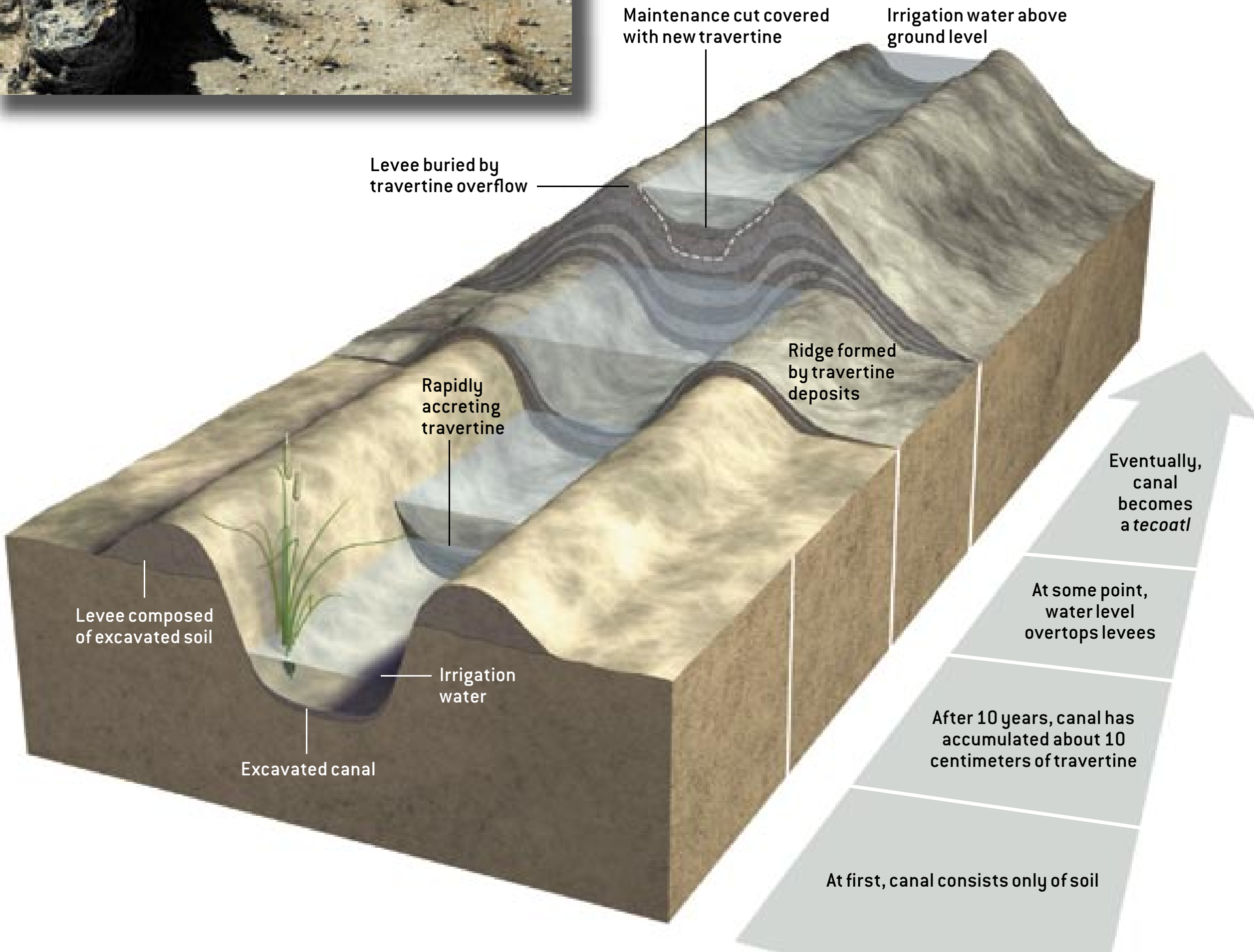
Where canals could not be constructed (across especially steep slopes, for example), the irrigators devised aboveground channels—aqueducts—built of carefully laid unmortared stone and compacted earth. Compared with the largest Roman aqueducts from roughly the same period, these were relatively simple but were nonetheless effective structures.

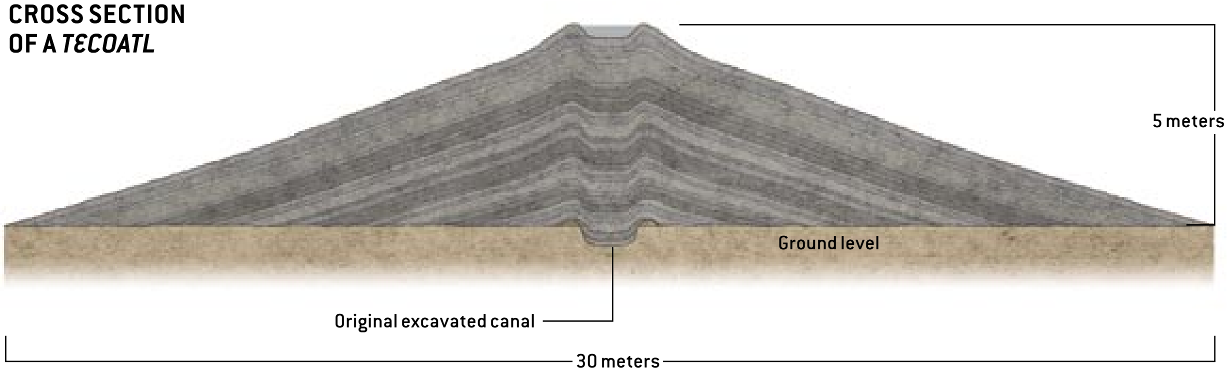
The Rio Xiquila Canyon provides an interesting example. This site has two aqueducts, at different levels above the river. Both were about one meter wide and had a near constant grade despite the irregular canyon walls. Fragments of pottery of known type and antiquity indicate the age of the structures. The one-kilometer-long lower aqueduct, constructed in about A.D. 400, lay only four to 12 meters above the river. This segment was vulnerable to damage from floods and landslides and was abandoned in 700. The upper aqueduct, probably built about that time, was 20 to 22 meters above the river and more than six kilometers long. It remained in use until at least 1540. The aqueducts carried relatively fresh river water and did not become fossilized.

Those canals that were fossilized provide an indelible history of their use and the environment in which they oper-

THE FORMATION OF A *TECOATL*

The stone snake, or *tecoatl,* in the photograph is one of thousands that crisscross the Tehuacán Valley. Each *tecoatl* originated as a canal excavated into the soil (*diagram*). As springwater flowed through the canal, mineral deposits precipitated from the water, eventually forming a crust so thick that the water level in the canal rose aboveground. Despite this change in grade, the *tecoatles* continued to operate, rising as much as five meters, widening to 30 meters at the base and forming sinuous ridges as long as 15 kilometers.





COURTESY OF JAMES A. NEELY (*photograph*); © MATTHEW FREY *Wood Ronsaville Harlin, Inc.* (*illustrations*)

ated. Travertine layers trapped abundant remains of aquatic algae, diatoms, mosses and larger marsh plants that grew in and along the canals. Because these organisms are sensitive to variations in environmental conditions, we can reconstruct the water's chemical composition, nutrient load, flow rate and clarity. Travertine also preserved pollen from plants cultivated in nearby fields, revealing that the canals irrigated corn, peppers and tomatoes. Amaranth, harvested in modern and quite possibly ancient times, was either planted or grew wild at the borders of the fields. Cattails thrived along the canals—probably without cultivation—just as they do today, and ancient farmers may have collected these useful plants for food, fiber and arrow shafts.

Because the canals created an artificial aquatic habitat unlike the surrounding semiarid environment, we can be certain that the aquatic plants that make up most of the organic remains in the travertine were absolutely contemporary with the period of canal use. The preserved organic matter gave us a remarkable opportunity: we could use radiocarbon dating, which is based on the chemical properties of carbon from living organisms, to measure the age of the canals directly. Ordinarily, the age of ancient water systems must be inferred from habitation sites and artifacts in their vicinity, an approach fraught with ambiguity, because it is not always evident that these chronological indicators were contemporary. Radiocarbon analyses solved this problem, demonstrating that some of the canals were built as long ago as 800 B.C. and that work on these structures continued until the early 1500s.

Terrace Irrigation

SOME 170 KILOMETERS southeast of the Tehuacán Valley, in the mountains at the southern end of the Valley of Oaxaca, a truly remarkable archaeological site illustrates the inventiveness and adaptability of Mexico's prehistoric hydraulic engineers. At Hierve el Agua, irrigation supported continuous farming for at least 18 centuries, from about 500 B.C. to 1350. Several large perennial springs with unusual characteristics furnished the irrigation water. *Hierve el agua* is Spanish for "the water boils," but the water is not hot. Instead it is naturally carbonated, like the groundwater sustaining the famous bottled water industry of southern France. Naturally carbonated water contains an abnormally large amount of dissolved carbon dioxide gas, derived from magma, from the metamorphosis of limestone, or from other complex processes. Deep underground, the water is confined under great pressure, which keeps the gas in solution in the same way that a container keeps carbonated beverages pressurized.

At Hierve el Agua, fractures in the rock provide escape routes, allowing groundwater to rise to the surface very rapidly. As the water emerges, the sudden release of pressure is like removing the cap from a shaken bottle of soda: the water effervesces spontaneously, releasing gas bubbles that create small geysers and cause the springwater to churn as if boiling. The dissolved carbon dioxide also makes the groundwater acidic, so much so that it dissolves the predominantly limestone bedrock. Like the travertine in the Tehuacán Valley, limestone is composed of calcite, and the springwater at Hierve el Agua contains extremely high concentrations of dissolved calcium and bicarbonate. Thus, layers of water-deposited travertine preserved archaeological records at Hierve el Agua much as they did in the Tehuacán Valley.

A place where cool water boils must have sparked the curiosity of the area's early inhabitants, who discovered that they could use the water for irrigation during months when the monsoons were not bringing rain. The steep slopes just below the spring lacked natural soil cover, however, so the farmers moved five million cubic meters of soil to the site by hand to develop nearly two square kilometers of terraced fields. They appear to have selected the soil carefully, possibly even sieving it to produce an even, porous texture, thereby improving drainage. They began constructing the terraces by dry-laying stone retaining walls at carefully spaced intervals across the bare slopes. They placed soil on the upslope side of each wall to create a narrow terrace that was level with the top of the wall. The farmers then fabricated a small

THREE PARALLEL *TECOATLES* (*right*): The smaller ones are branches that led to nearby ancient fields. Beyond the white wall in the middle ground, the unexcavated main *tecoatl* can be traced toward what is now the center of Tehuacán City, where modern street construction (*below*) has damaged parts of the canal.

COURTESY OF JAMES A. NEELY

"wall canal" atop each wall, in all constructing and maintaining more than 6.5 kilometers of such canals [*see illustration on pages 78 and 79*].

The wall canals had a gentle inclination, allowing water to be shunted into them from much larger water supply canals running directly down the slope from the springs. Additional small canals connected the downstream ends of the wall canals and carried water to terraces lower on the slope or returned flow to the supply canals. As water passed through a wall canal, it pooled in shallow circular basins, or *pocitos,* placed every few meters along its length in the floor of the canal. By dipping small vessels into the *pocitos,* the farmers obtained water for hand irrigating the plants growing on either adjacent terrace. In Spanish, this type of hand watering is known as *riego a brazo,* and it is still practiced in the region.

Close inspection of the terraces and canals reveals the brilliance of their design. The regular wall spacing and narrow width of the terraces restricted the amount of soil needed to create each terrace, while the immediate proximity of the wall canals and *pocitos* ensured ease of watering. Even on the steepest slopes, where the height of the retaining walls was as much as 2.4 meters, the width of the terraces was held relatively constant. Most are 2.4 to three meters wide, equivalent to two comfortable arm spans of a person 1.4 to 1.7 meters tall, which, as skeletal remains show, was the range in height of the ancient farmers. Small "weep holes" at the base of each terrace wall enhanced drainage and recapture of soil moisture. Hand watering and proper soil drainage were important because of the high mineral content of the water. If the amount of water were excessive or the water failed to drain properly, minerals would build up quickly, making the soil nonporous and too hard to turn by hand or to allow root growth.

Flow was continually rerouted through the entire canal network, conveying water only to those wall canals where it was needed at a given time. Consequently, no single wall canal carried enough water to become a large

PURRÓN DAM (*at lower right in photograph above*) near Puebla, Mexico, is believed to be the largest water retention structure built in the Americas before the arrival of the Europeans. Construction began in 750 B.C. and continued at intervals until about A.D. 1150, when the dam stood almost 25 meters high and stretched 400 meters in length. The eroded southern profile of the dam (*far right*) dwarfs a six-foot-tall man.

tecoatl. Instead a relatively thin layer of travertine coats these canals, preserving many of the construction details. Particularly interesting is the absence of sluices or gaps through which water might have been diverted from the canals onto the terraces in quantity. This type of irrigation, known as flooding, would have caused the terraces to become entirely encrusted or infused with travertine, probably after only a few applications. By adopting a highly efficient method of hand watering, the farmers reduced mineral accumulation in the soil while minimizing the amount of water needed to maintain the plants on each terrace, thus increasing the total area that could be irrigated.

Watering was largely confined to the dry season. During the remainder of the year, rainwater helped to flush the accumulated minerals from the porous soil. This process was enhanced by decomposition of organic matter within the terrace soil. In addition to unharvested crop stubble, the organic matter may have included night soil and other domestic wastes, which were mixed in routinely to renew fertility.

We also found evidence for soil amendment using household debris: fragments of pottery of different ages lie within the terrace soil in chronological order, from the bottom to the top of the fill. Ceramic vessels can sometimes be dated directly by radiocarbon techniques or indirectly through stylistic classification of their form, composition and color pattern, once a particular type has been found at a site with an established chronology. Besides providing a record of ongoing waste disposal, the pottery at Hierve el Agua had the unintended archaeological benefit of defining which kinds of vessels were in everyday use and therefore most likely to be broken. In this way, we learned that the people who were working the fields disposed of trash that included both everyday domestic ceramics and finer serving wares, whereas much finer vessels are found only at a small temple at this site.

THE AUTHORS

S. CHRISTOPHER CARAN and *JAMES A. NEELY* share a longtime interest in the prehistoric water management systems of Mexico and the American Southwest. Caran has been a research geologist at the University of Texas at Austin, specializing in Quaternary studies, and is currently president of Quaternary Analysis Laboratories in Austin. Neely is professor emeritus in the department of anthropology at the University of Texas at Austin; his research has focused on the development of agriculture. Neely discovered most of the water management systems discussed in this article in the 1960s and 1970s; he has investigated them intensively since then, in part with Caran from 1988.

COURTESY OF JAMES A. NEELY

COURTESY OF JAMES A. NEELY

A Technological Saga

THE SEEMINGLY ABRUPT origin of fully developed, large-scale irrigation technology appears puzzling at first. But this apparent absence of more modest forerunners is most likely the result of gaps in the archaeological record. Our discovery in 1993 of what may be the oldest water well in the New World indicates that water management may have had a much older, albeit embryonic, beginning than was previously recognized. Excavated nearly 10,000 years ago, the well was five meters deep and 10 meters in diameter at the former ground level, which was subsequently buried. It may have remained in use for 2,000 years. The well, in what is now the village of San Marcos Necoxtla in the Tehuacán Valley, probably predates New World agriculture. Although it was most likely not used for irrigation, it does provide evidence that water management in this region began very early indeed.

We have found no examples of hydraulic construction in the intervening centuries between the digging of this well and about 3,000 years ago, when the first canals appear. It is, however, probable that small wells, temporary weirs for diverting water from streams, or other simple water supply features were built during this period. The earliest cultivation may have initially required water hauling or small-scale irrigation using canals that were not preserved or have not been discovered.

But the question still remains of how the early hydraulic engineers of Mexico managed to lay out canal routes many kilometers long over highly irregular terrain, while maintaining a continuous downward gradient of less than two degrees. Today comparable construction would be impossible without sophisticated surveying instruments. The ancient Egyptians used levels and calibrated rods for sighting over long distances. Although such simple yet effective tools and methods must also have been available to the engineers in Mexico, we have no direct evidence of these details.

We do have a partial answer to how they planned the canal systems. At a remote locality in the Tehuacán Valley, a line of small boulders leads away from a sharp bend in a *tecoatl*. This line extends down the short slope of a notch in a ridge, then up the other side to a slightly higher point immediately above a small valley lacking canals. The line of boulders may have been a "blueprint" for future canal construction. For water to reach the other side of the notch, the existing *tecoatl* would have to build upward at least a meter before the new branch could be constructed. If the builders allowed for normal travertine accretion, that goal might be reached in a century, at which time the farmers' descendants could add another irrigated field to the system.

Whether these projects were developed and controlled by individual users or by a more centralized authority is another of the many questions fascinating archaeologists. In the 1950s the well-known historian Karl A. Wittfogel advanced the hypothesis that large-scale exploitation and distribution of water resources were essential steps in the rise of civilizations worldwide. According to this principle, only "hydraulic societies" achieved the hallmarks of a sophisticated culture, such as permanent agriculture, economic diversity, record keeping and hierarchical administration. A hydraulic society became civilized because a reliable water source provided both the incentive and capacity to do so. Yet the inverse also seems implicit: construction and maintenance of an extensive water management infrastructure might require the focused attention of a well-organized state. Other investigators have questioned both these propositions, pointing out that small, loosely organized sociopolitical entities could build and operate water systems of at least moderate scale, perhaps in cooperation with similar neighboring organizations but without a central authority.

Evidence exists for each of these interpretations. Modern irrigation in the Tehuacán Valley, for example, is managed by community-based, nongovernmental *sociedades de agua* ("water societies"), which trace their ancestry to native traditions. Even today these highly coveted water rights are often conveyed by inheritance, a practice that can be traced back to precolonial times in the Aztec codices and in early Spanish documents. Each small community is responsible for proper use and maintenance of its part of a larger canal system, but overall management is effected by consensus among the various partner communities. Thus, the system operates both locally and collectively.

The debate about how ancient societies constructed and managed their hydraulic infrastructure will continue. What is not controversial is that the systems in southern Mexico stand as engineering marvels, ranking among the pinnacle achievements of prehistoric builders anywhere in the world. SA

MORE TO EXPLORE

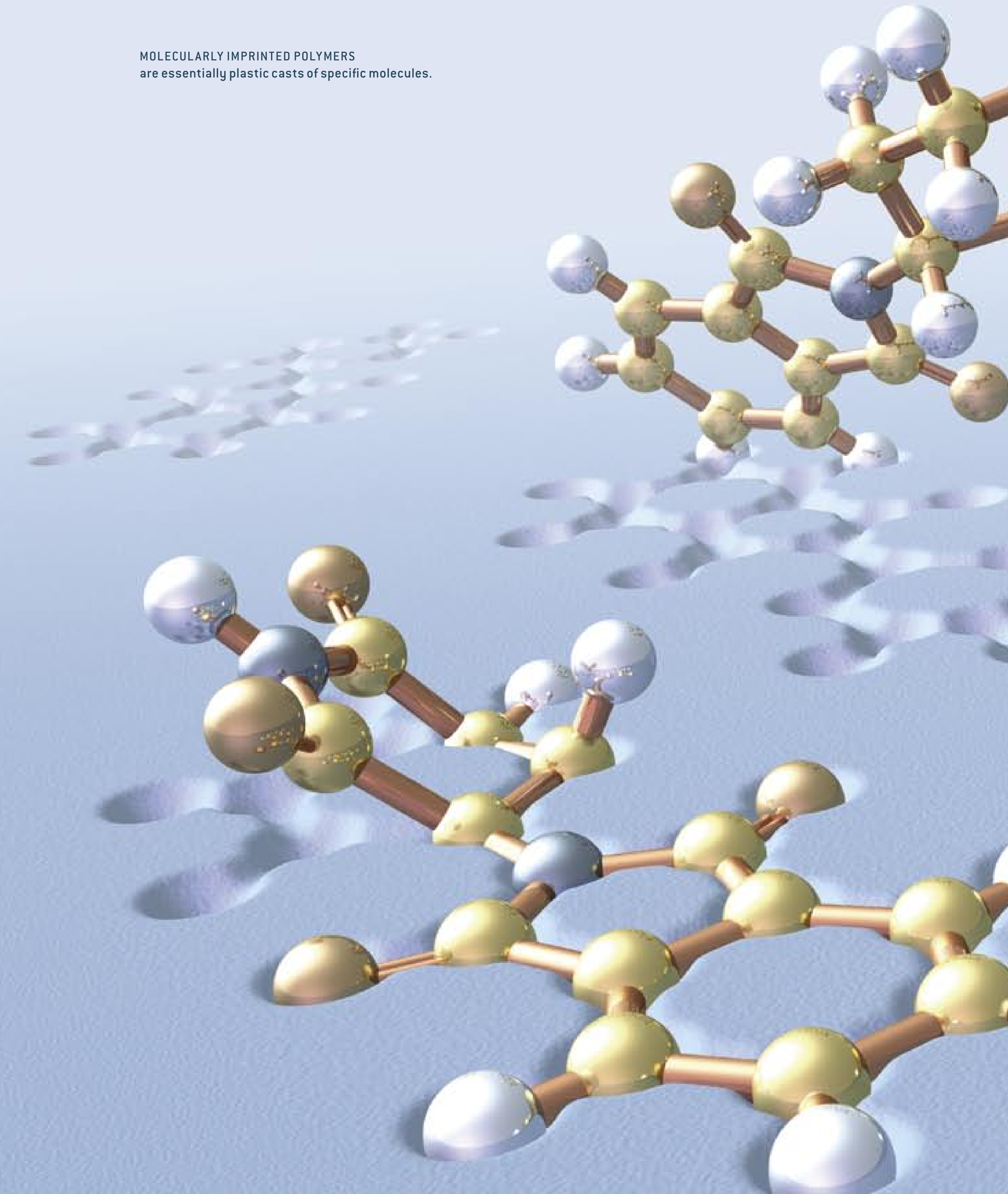
The Keepers of Water and Earth: Mexican Rural Social Organization and Irrigation. K. I. Enge and Scott Whiteford. University of Texas Press, 1989.

Irrigated Agriculture at Hierve el Agua, Oaxaca, Mexico. J. A. Neely et al. in *Debating Oaxaca Archaeology.* Edited by J. Marcus. University of Michigan, Museum of Anthropology, Anthropological Papers No. 84, 1990.

A Late Paleo-Indian/Early Archaic Water Well in Mexico: Possible Oldest Water-Management Feature in the New World. S. C. Caran et al. in *Geoarchaeology: An International Journal,* Vol. 11, No. 1, pages 1–36; January 1996.

A Contextual Study of the "Fossilized" Prehispanic Canal Systems of the Tehuacán Valley, Puebla, Mexico. J. A. Neely in *Antiquity,* Vol. 75, No. 289, pages 505–506; 2002.

MOLECULARLY IMPRINTED POLYMERS
are essentially plastic casts of specific molecules.



The Promise of MOLECULAR IMPRINTING

Tiny plastic imprints and mimics of biological molecules are poised to speed drug discovery, warn of bioterror attacks and remove toxins from the environment, among other applications

By Klaus Mosbach

More than three decades ago my students and I at Lund University in Sweden, along with other teams, developed "fishing nets" of sorts that worked at the nanometer scale. The nets we created could trap living cells and, later, smaller biological entities, such as enzymes or other molecules. Under the right conditions, our "catches" could continue to do their usual tasks outside of living organisms for months.

The technology proved attractive for dozens of applications [see "Enzymes Bound to Artificial Matrixes," by Klaus Mosbach; Scientific American, March 1971]. For instance, plastic nets containing *Escherichia coli* cells are used today to produce aspartic acid, an amino acid used in the preparation of various medicines. In the food industry, plastic embedded with a specific enzyme converts the sugar glucose into fructose, which is much sweeter. A different net-and-enzyme combination has even helped to fabricate the precursors of the plastic material that makes up the nets. And, to our delight, potential applications for the traps keep arising, including in medicine. Notably, cells held in the nets might replace ones that have died or malfunctioned, such as the insulin-producing cells that are needed by diabetics.

Among the potentially HARMFUL substances that MIPs have detected in tests is the herbicide ATRAZINE.

But the original netting tool represents just a first attempt at devising technologies that marry plastics and molecules. Today more than 500 researchers around the world are testing applications of a second technology, molecular imprinting, which I played a critical role in developing. My group devised a now popular version using principles of biochemistry. Other investigators—notably Günter Wulff of Heinrich-Heine University in Düsseldorf, Germany, and Kenneth J. Shea of the University of California, Irvine—developed methods based on principles of organic chemistry.

In general, plastic beads or other structures are covered with imprints of specific molecules, and the imprints—really, casts of the molecules—are put to use for selected jobs. Once perfected, this technology could have applications in many areas, including the food industry, which sees benefit in its ability to remove contaminants, such as the fungal poison aflatoxin. The new tools have also begun to draw much interest from the biomedical field, where they could speed the early stages of drug discovery (bringing down costs), be used for drug purification and separation, and contribute to the development of medical devices and diagnostic tools.

Moonshine Research

AS MY LABORATORY at Lund was studying methods to immobilize enzymes and cells, I started to ask what might happen if we modified our original "fishing" technology so that the netting molded itself around the molecules it captured in a way that let us flush the catches from the nets, leaving permanent cavities, or imprints. Would these imprints enable other molecules of the same type as the originals to fit into those holes? If so, I realized, the imprints might be useful for various tasks, including isolating selected molecules from mixtures of compounds (because only those molecules having the right shape and chemical groups would settle snugly into the cavities).

For more than 20 years, my team intermittently carried out experiments to develop this second technology. We did what I call "moonshine research," because we had to prove the principle behind the idea on our own, without formal funding. During that time, I heard unofficially that potential financiers thought my notion seemed too far-fetched, too much like magic.

Over the decades we figured out ways to make plastic imprints of molecules, and outside funds started coming in. We also managed to keep the procedure simple; it takes only a few days to make beads or thin films with hundreds of thousands of imprints. First, a technician mixes molecules of interest—what we call templates—with selected building blocks of plastics. These units, also called monomers, create a plastic web around each template. A solvent then dissolves the templates, leaving a material mottled with plastic-coated cavities that have a "memory" for the contours and charge distributions of chemical groups in the original biological molecule.

The plastic imprints, named molecularly imprinted polymers—or, as I dubbed them, MIPs—have several attractive features. They are inexpensive, because they require relatively little time to produce and because the plastic monomers are cheap. And, much like their forebears, they can be stable for a long time, even under extreme conditions. Some of the preparations in our lab have remained functional for as long as a year.

One of the many potential uses for MIPs is the removal of unwanted substances from blood. Imprints of the substances in question would be displayed on plastic beads packed in tubes. A patient with, say, failed kidneys could rely on one tube, or multiple tubes gathered together, in a device kept outside his or her body to clear a dangerous substance from the blood system. As the person's blood passed through an intravenous tube running between a vein and the exterior MIP tubing, the beads would collect the selected substance, and the cleansed fluid would then reenter the circulation. Used at all times, such treatment could, in theory, reduce the frequency of needed hemodialysis. On being filled with the unwanted matter, the MIP unit could be removed and replaced by another one.

Imprinted devices might eventually be designed to pull unwanted substances from other body parts as well, such as the gastrointestinal tract. Imprints of cholesterol molecules, for example, can extract cholesterol from solutions.

The selectivity of MIPs could also be exploited by the pharmaceutical industry to make purer drugs. Such purity can become especially important when a drug molecule occurs in two mirror-image forms, one of which is helpful and the other potentially harmful. Thalidomide is the classic example. Before anyone realized that the helpful form had an injurious counterpart, formula-

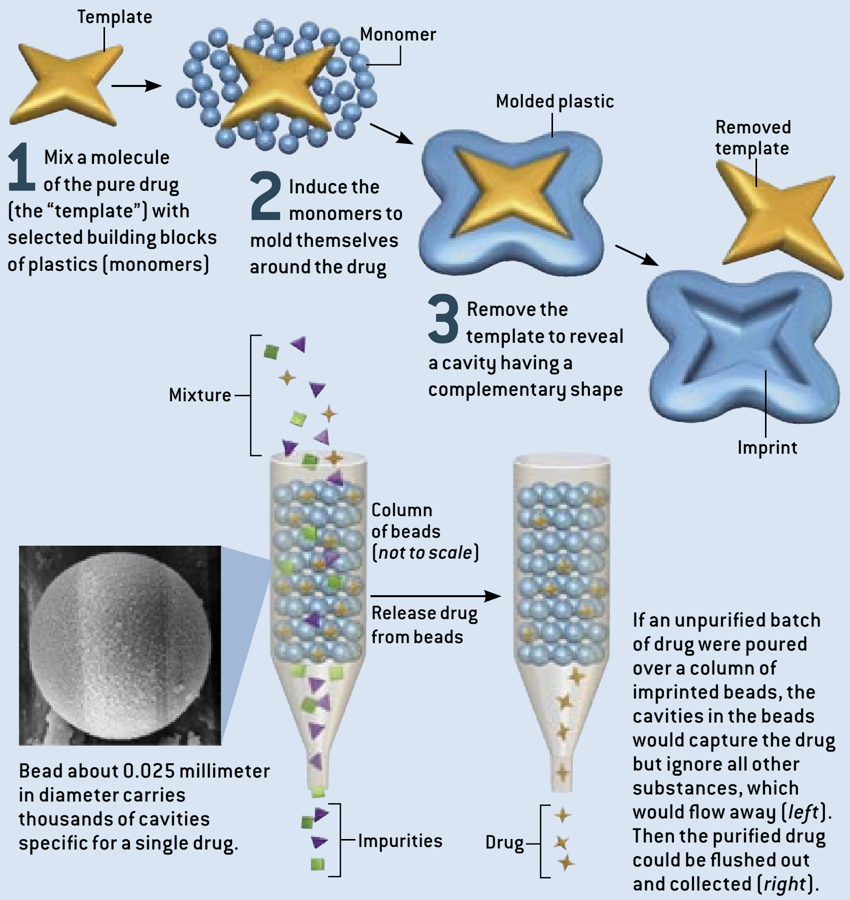
Overview/*Nanoplastic Applications*

- Plastic imprints of specific molecules—molecularly imprinted polymers (MIPs)—will capture those molecules and only those molecules. MIPs can thus be used to separate desired substances from an impure mixture or to detect pathogens or toxins in the environment or in blood samples.
- Start-up companies are now working to commercialize the technology.
- Next-generation variations on the technology, known as double imprinting and direct molding, are also under development.

KEN EWARD *BioGrafx* (*preceding pages*)

HOW MIPS ARE MADE

Molecularly imprinted polymers, or MIPs, have many uses, such as removing impurities from a batch of a newly made pharmaceutical (*bottom*). To make a MIP for this purpose, technicians would essentially follow steps 1–3 below.



GEORGE RETSECK; COURTESY OF KLAUS MOSBACH AND ANDREW MAYES (*micrograph*)

tions containing both versions were prescribed to women for a variety of reasons in dozens of countries during the late 1950s and early 1960s. Tragically, the "bad" form of the molecule led to severe malformations in perhaps 10,000 babies born to women who took the medicine during pregnancy.

Drug manufacturers often synthesize separate forms of mirror-image (or "right-handed" and "left-handed") molecules, but the production methods can result in formation of small amounts of the unwanted molecular form. Compared with typical commercial techniques, those relying on MIPs might prove more efficient for detecting and removing the unwanted incarnation of a drug molecule, because each form would fit well only into its corresponding cavity.

Given the detection abilities of MIPs, some companies and governmental bodies concerned with terrorism and emerging diseases are eyeing them as components in sensors for sniffing out toxins and pathogens (disease-causing agents). Although sensors can be outfitted today with biological molecules that do the "sniffing," sometimes those are not robust enough to survive in more rugged environments than laboratory confines.

Among the potentially harmful substances that MIPs have detected in tests is the herbicide atrazine. The plastic cavities have also identified sarin, a nerve gas that can be used as a bioterrorism weapon. (The release of sarin for this purpose became well known when a religious cult twice released the toxin in Japan in the mid-1990s. The gas killed 19 people and injured thousands.) MIPs could be exploited as well to sense anthrax spores, the notorious agents that were stuffed into envelopes and sent to U.S. government officials and members of the media in the fall of 2001.

Still more enticing, a single sensor unit equipped with multiple kinds of MIPs could identify several compounds in an individual sample at once. Such a multifunctional sensor could be constructed on an electronic chip that, on detecting one of the target substances, would send "word" of its presence to a receiver. Separate containers with large numbers of MIPs could then remove the undesirable materials. Government and industry have indicated considerable interest in units along those lines for the environmental cleanup of lakes, creeks and soil.

Form Equals Function

WHEREAS MANY APPLICATIONS of MIPs rely on their ability to capture the molecules or microorganisms that match a particular template, other applications involve molding plastic to mimic all or part of a natural molecule, such as an antibody. The body makes antibodies naturally when the immune system notices specific components, or antigens, of a foreign substance, such as a virus or bacterium that has found a way inside a living host. Antibody molecules have great specificity—each type binds tightly to a particular molecule while ignoring all others, like a key fits into a lock. For that reason, the developers of diagnostic tests have long exploited them. When exposed to a blood sample, for

THE AUTHOR

KLAUS MOSBACH is professor and founder of the department of pure and applied biochemistry and the Center of Molecular Imprinting at the Lund University in Sweden. He was also cofounder of the department of biotechnology at the Swiss Federal Institute of Technology (ETH) in Zurich and a founder with others of MIP-Globe, a company focused on molecular imprinting. A trained pianist, he derives his scientific inspirations from playing music.

We are ENCOURAGED by growing interest from the pharmaceutical and biotechnology INDUSTRIES.

instance, certain antibodies will bind to a particular bacterium if it is present, indicating that a person is infected. Other antibodies might reveal the concentrations of various proteins in the blood.

One way that makers of diagnostic tools generate needed quantities of antibodies is by injecting foreign proteins or other compounds into goats or other animals. Technicians then bleed the animals for the resulting antibodies. But long-lasting mimics of antibodies can also be made by imprinting plastic with a specific antigen; the resulting MIPs then possess essentially the same antigen binding site as the corresponding antibody has. Such "plastibodies," as my lab calls them, could potentially take the place of natural antibodies in many tests, thereby reducing the need for involving animals. (Incidentally, the development of plastibodies, for which I received the Nordic Prize for Alternatives to Animal Experiments, was the only part of my scientific work that one of my daughters fully understood and applauded when she was a teenager.)

MIPs might also serve as long-acting substitutes for enzymes in industry. In nature, every organism produces thousands of enzymes, each of which catalyzes a specific biochemical reaction, such as cleaving a particular molecule in a set place or fusing two substances. Normally, the reaction occurs when the enzyme's target, or substrate, fits into a groove on the enzyme known as the active site.

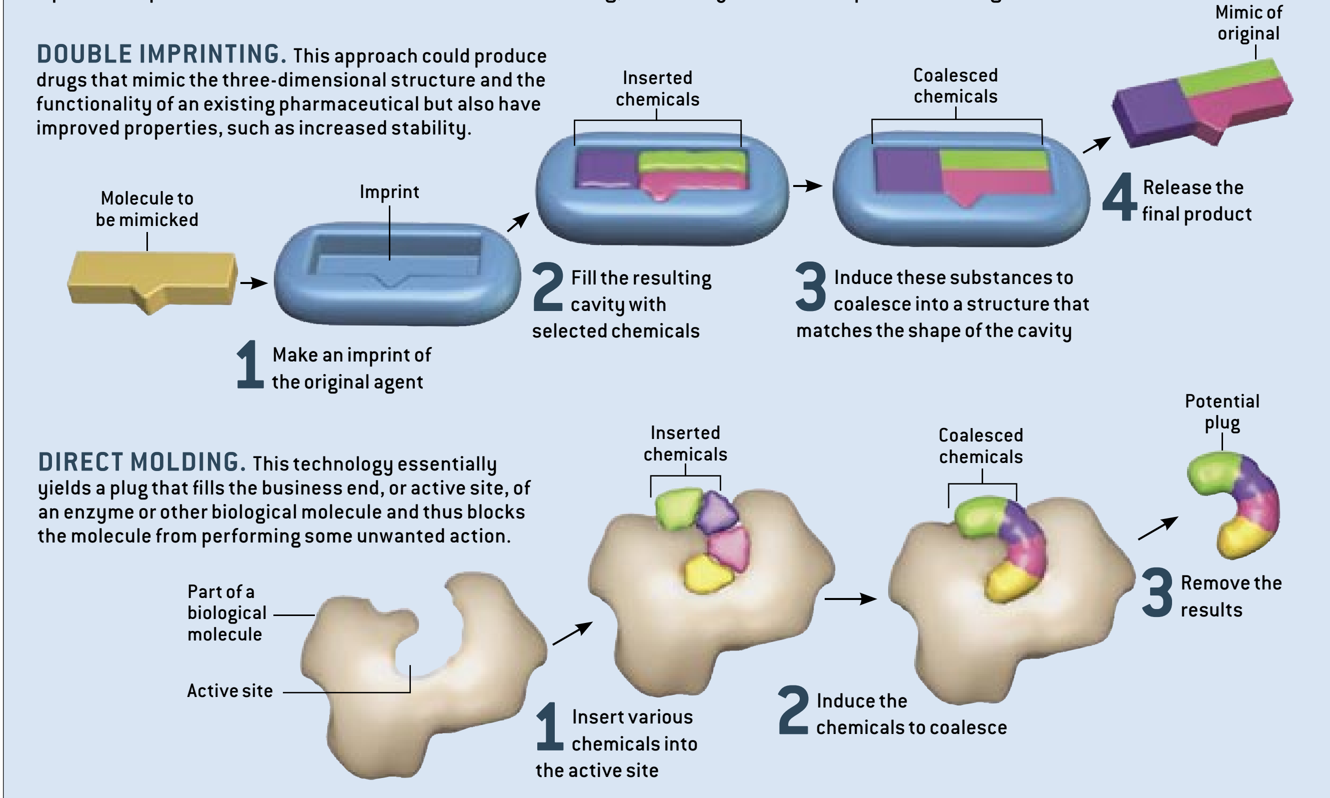
To make artificial enzymes, or "plastizymes," my lab and other groups are attempting to create plastic cavities that are imprints of particular conformations of substrates and that thus mimic the three-dimensional shape of the active site of a real enzyme. The monomers we rely on, and the resulting plastizymes, have chemical groups similar to those occurring in natural enzymes. The first efforts along these lines have resulted in some enzymatic activity, but we still have to find ways to make the MIPs operate more effectively. Plastizymes could conceivably perform actions that no natural enzyme has yet been discovered to carry out—such as detoxifying certain substances by breaking them apart.

The Next Generation

MY GROUP is now pursuing two offshoots of MIP technology. One of these processes yields a mimic of an original molecule. More precisely, it generates a

BEYOND MIPS: MAKING POSITIVES FROM NEGATIVES

Two new technologies exploit cavities as molds for making useful compounds. In double imprinting, the mold is artificial: a plastic imprint of a selected substance. In direct molding, the cavity is natural: a part of a biological molecule.



GEORGE RETSECK

substance having the same three-dimensional shape and functional abilities as the original—the functionality deriving both from the matching shape and from having the same arrangement of charges at defined positions. We refer to this technology as double imprinting because it involves making a new molecule from an imprint—essentially, an imprint of an imprint. After making the first impression, we use the resulting cavity as a tiny mold, or nanovessel, and place fragments of molecules or precursors of plastic polymers into that nanovessel. Subsequently we allow the components to combine into a single structure matching the shape of the small cavity.

This strategy might aid pharmaceutical firms that synthesize medicines related to ones already available. Drug companies make such relatives to improve on the originals or to produce compounds serving the same purpose as ones patented by competitors. But today's typical method requires developing what is known as a combinatorial library, which can consist of tens of thousands of related compounds. Each individual compound is tested for its ability to bind to a given biological molecule (which, incidentally, could be replaced by a more stable MIP counterpart). Then the substances that bind are tested further. Double imprinting is more focused. We mix and match materials poured into a nanovessel created by imprinting, looking to see which combinations have promising properties. Compared with the screening of combinatorial libraries, our procedure could dramatically limit the number of steps needed to find good candidates for follow-up testing.

The other technology, which we call direct molding, is being worked on by a few groups. It employs a biological molecule (often an enzyme or the part containing its active site) as something of a nanovessel for concocting a new drug—an approach that, for instance, could speed the discovery of agents that inhibit selected enzymes. Roughly a third of the medicines on the market today are enzyme inhibitors, and more efficient methods for discovering other members of this category would be valuable.

Let us say a drugmaker wanted to shut off an enzyme that catalyzes interactions involved in the metastatic growth of tumors. A good solution would be to create a molecule that plugged the enzyme's active site, preventing the site from interacting with its usual substrate. Investigators could blindly screen all kinds of compounds, hoping to hit on a great plug. Or they could insert monomers and other small chemicals into the active site—much as is done in the double-imprinting method. A combination that resulted in a tight-fitting unit could then be tested to see if it inhibited the enzyme in a living cell.

As with all new technologies and specific applications, certain manufacturing problems related to MIPs have to be addressed. We need to figure out how to scale up the quantities of the imprints we can make. We must ensure that every copy of an imprint is always the same as the others. And we want to develop efficient ways to flush out the templates.

As those of us in this burgeoning field grapple with the remaining obstacles for improving MIPs and their successors, we are encouraged by growing interest from the pharmaceutical and biotechnology industries. I am continually amazed by the realization that people today can fabricate in days the kinds of molecular shapes that nature took millions of years to develop. I look forward to the time when this ability is in widespread use to hasten drug discovery and to support an array of other applications. SA

Some Companies Using MIPs

COMPANY	SAMPLE ACTIVITIES
Aspira Biosystems Burlingame, Calif.	Employs a selected segment of a protein as a template in a technique called partial molecular imprinting; expects the approach to be applied in medical research and therapy
MIP-Globe Zurich, Switzerland	Pursues most aspects of molecular imprinting, including applications related to drug discovery
MIP*Solutions* Las Vegas, Nev.	Develops technology to help provide safe drinking water through removing contaminants and for use in wastewater treatment and water-based mining operations
MIP Technologies Lund, Sweden	Designs tools for extraction and separation of substances from complex mixtures at analytic and industrial scales; the technology is intended for pharmaceutical, chemical, food and other industries
POLYIntell Rouen, France	Designs polymers for use in purification and sensing; generates artificial antibodies and enzymes on demand, for use in pharmaceutical and other industries
Semorex North Brunswick, N.J., and Ness Ziona, Israel	Develops handheld devices that help physicians to diagnose infectious diseases and early cancers or that allow military, security and emergency workers to rapidly detect and identify chemical warfare agents and explosives in the field

MORE TO EXPLORE

Drug Assay Using Antibody Mimics Made by Molecular Imprinting. G. Vlatakis, L. I. Andersson, R. Müller and K. Mosbach in *Nature,* Vol. 361, pages 645–647; February 18, 1993.

The Emerging Technique of Molecular Imprinting and Its Future Impact on Biotechnology. K. Mosbach and O. Ramström in *Bio/Technology,* Vol. 14, pages 163–170; February 1996.

Formation of a Class of Enzyme Inhibitors (Drugs) Including a Chiral Compound by Using Imprinted Polymers or Biomolecules as Molecular-Scale Reaction Vessels. Y. Yu, L. Ye, K. Haupt and K. Mosbach in *Angewandte Chemie: International Edition,* Vol. 41, pages 4459–4463; 2002.

Two Ways to Shape New Drugs. S. Borman in *Chemical and Engineering News,* Vol. 81, No. 2, page 40; 2003.

Molecularly Imprinted Materials Science and Technology. Edited by M. Yan and O. Ramström. CRC Press, 2004.

Klaus Mosbach's Web sites: **www.klausmosbach.com** and **www.MIP-Globe.com**

WORKINGKNOWLEDGE

IMAGE STABILIZATION

Steady Cam

"Hold still while I take your picture."

Photographers who instruct their subjects this way can still end up with a blurry image if they themselves move their hands even slightly when depressing the shutter, shaking the camera. The problem is so common among people who use digital cameras and camcorders, especially in low light when the shutter must stay open longer, that manufacturers are introducing image stabilization systems that automatically correct for human shudder. "The industry is moving toward cameras with higher megapixels, smaller size and longer zoom lenses that magnify shake," says Jay Endsley, manager of digital camera advanced development at Eastman Kodak Company in Rochester, N.Y. "So we're adding whatever we can to improve picture quality."

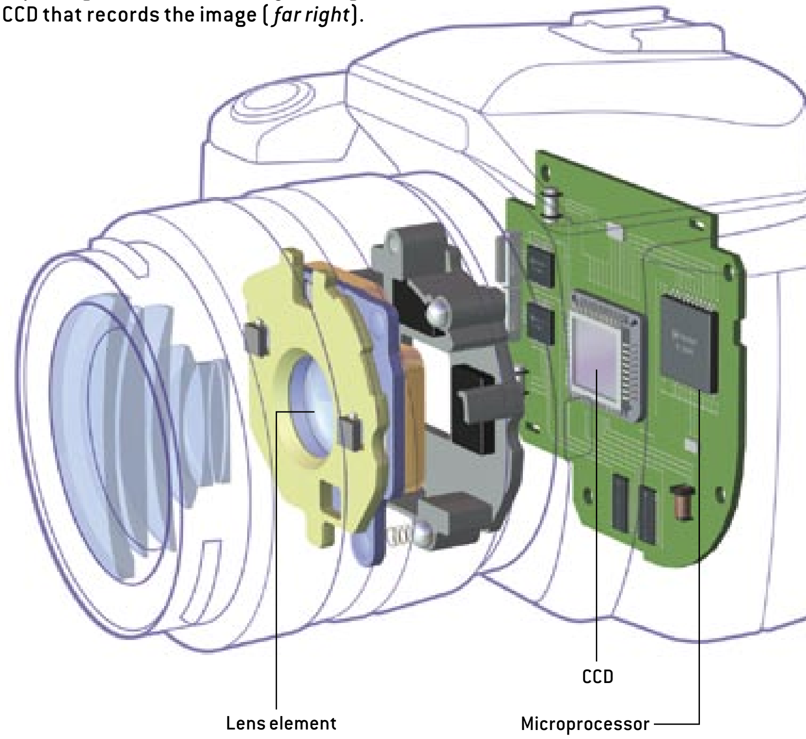
Digital cameras employ two different image stabilization hardware schemes. One system moves a segment of the lens to deflect incoming light, compensating for the direction of shake; the other moves the CCD—the sensor that captures the image [*see top illustrations*]. Engineers who favor the lens approach say moving the CCD complicates the recording of sharp images, especially when using a flash. Designers of the CCD method note that it works with every lens a user might attach to the camera, negating the need to buy different lenses.

Correction inside camcorders also occurs in one of two ways: by moving a bellows inside the lens assembly or by digitally adjusting which CCD pixels record the incoming light [*see bottom illustrations*]. The latter method, sometimes called electronic image stabilization, is simpler and cheaper to implement but reduces the camcorder's field of view by 10 to 20 percent.

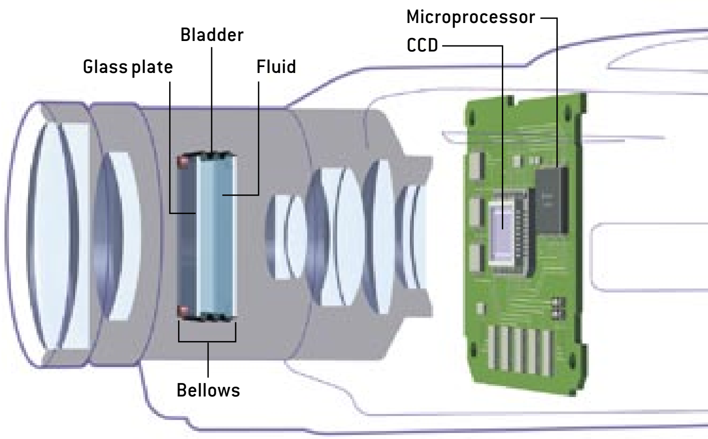
Developers are trying to devise digital processes for still cameras, but so far the approaches require much more processing power than current models carry. All the systems now deployed raise cost and add complexity. "Image stabilization began in high-end cameras because they had longer zooms and therefore needed greater shake compensation," says Chuck Westfall, who heads the technical information department at Canon USA in Lake Success, N.Y. "But 10× and even 12× optical zoom is moving down to less expensive cameras, so image stabilization must move down, too. It adds cost, but customers find it's worth it." —*Mark Fischetti*

GEORGE RETSECK; SOURCES: EASTMAN KODAK AND CANON USA; ENVISION/CORBIS (*photographs*)

DIGITAL CAMERA can offset blur from a shaky photographer's hand by rapidly adjusting a lens element or by moving the CCD that records the image (*far right*).



DIGITAL CAMCORDER contains a variable-angle prism, or bellows, that refracts light to offset shake. An alternative method, with no bellows, is digital stabilization (*far right*).



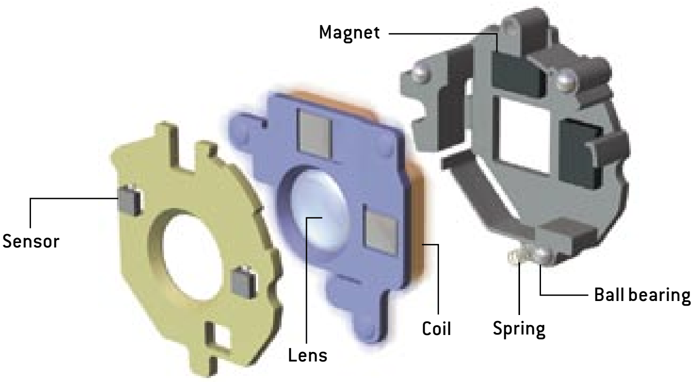
DID YOU KNOW . . .

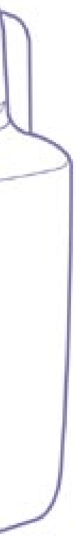
➤ PANNED: When shooting a movie, it is common to pan—move the camera to follow the action. But a stabilization system will try to counter that motion. Software can override the correction; if all the pixels are moving in harmony for more than a moment, the change is deemed to be a pan, not an accidental bump. No system can tell the difference during the first fractions of a second, however, so some cameras allow the operator to turn stabilization off.

➤ CELL PHONES: Image stabilization is hard to cram into cell phone cameras because there is so little space. "The phones also have to pass a drop test," Kodak's Jay Endsley notes, "which is tough for mechanical correction. Digital stabilization could work well, though."

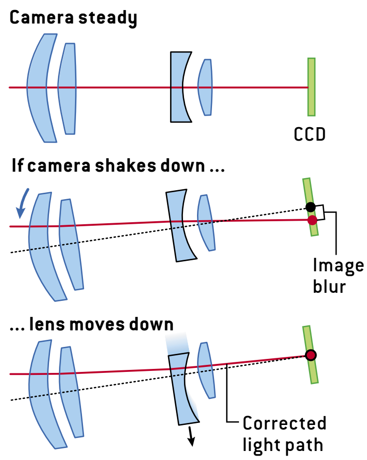
➤ BINOCULARS: Even the slightest hand shake will blur the image seen through strong binoculars, because the high magnification compounds the error. To compensate, Canon has inserted the same corrective bellows from camcorders into each cylinder of its high-end binoculars.

➤ SPEEDY ANALOGY: Brochures espousing cameras with image stabilization tend to say that the system corrects up to two or three shutter-speed stops. The idea is that if, without correction, a user could not hold still enough to take a blur-free shot at a shutter speed slower than $1/125$ of a second, with correction he or she would in effect be steady down to $1/60$, $1/30$ and possibly $1/15$ of a second.

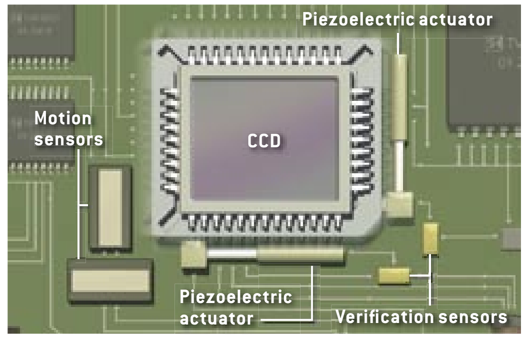


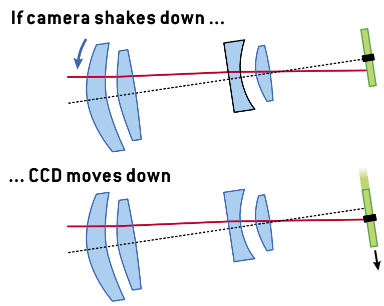


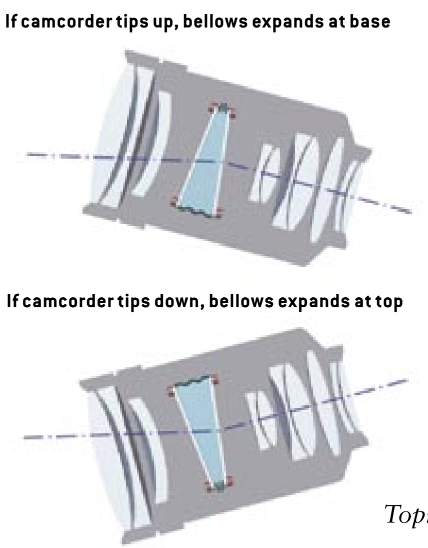
LENS ELEMENT shifts freely on ball bearings—up, down, left or right—inside a housing, when magnets induce attached coils to move. (Other designs nudge the lens with piezoelectric rods.) Sensors detect the direction and speed of camera jitter, signaling a microprocessor that controls the magnets. For example, if the camera tips down, the microprocessor instructs the lens element to drop, refracting the incoming light upward. A spring anchors the lens in a home position.



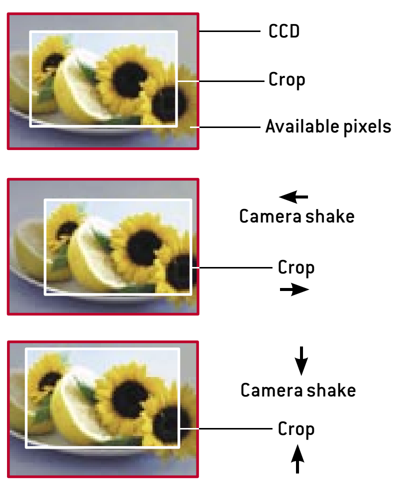
CCD is moved up, down, left or right by piezoelectric actuators when directed by a microprocessor, if motion sensors indicate the camera is jiggling. Verification sensors confirm that the CCD has shifted properly. If the camera tips down (*below*), the CCD drops so light stays centered.







DIGITAL STABILIZATION requires no moving parts. Software crops the incoming image so it covers only 80 to 90 percent of the CCD's pixels. If electronics sense that the camera shakes left, the software adjusts the crop to the right (over available pixels); if it shakes down, the crop moves up.



Topic from reader William Phillips. Send ideas to workingknowledge@sciam.com

Scientists on Religion

THEIST AND MATERIALIST PONDER THE PLACE OF HUMANITY IN THE UNIVERSE **BY GEORGE JOHNSON**

GOD'S UNIVERSE
by Owen Gingerich
Belknap Press (Harvard University Press), 2006 ($16.95)

THE LANGUAGE OF GOD: A SCIENTIST PRESENTS EVIDENCE FOR BELIEF
by Francis S. Collins
Free Press (Simon & Schuster), 2006 ($26)

THE GOD DELUSION
by Richard Dawkins
Houghton Mifflin, 2006 ($27)

THE VARIETIES OF SCIENTIFIC EXPERIENCE: A PERSONAL VIEW OF THE SEARCH FOR GOD
by Carl Sagan. Edited by Ann Druyan
Penguin Press, 2006 ($27.95)

Ten years after his death in 1996, science writer Walter Sullivan's byline occasionally still appears in the *New York Times* on obituaries of important physicists, as though he were beckoning them to some quantum-mechanical heaven. This is not a case of necromancy—the background material for *Times* obits is often written in advance and stored. If the dead really did communicate with the living, that would be a scientific event as monumental as the discovery of electromagnetic induction, radioactive decay or the expansion of the universe. Laboratories and observatories all over the world would be fiercely competing to understand a new phenomenon. One can imagine Mr. Sullivan, the ultimate foreign correspondent, eagerly reporting the story from the other side.

Light is carried by photons, gravity by gravitons. If there is such a thing as spiritual communication, there must be a means of conveyance: some kind of "spiritons"—ripples, perhaps, in one of M Theory's leftover dimensions. Some theologians might scoff at that remark, yet there has been a resurgence in recent years of "natural theology"—the attempt to justify religious teachings not through faith and scripture but through rational argument, astronomical observations and even experiments on the healing effects of prayer. The intent is to prove that, Carl Sagan be damned, we are not lost among billions and billions of stars in billions and billions of galaxies, that the universe was created and is sustained for the benefit of God's creatures, the inhabitants of the third rock from the sun.

In *God's Universe,* Owen Gingerich, a Harvard University astronomer and science historian, tells how in the 1980s he was part of an effort to produce a kind of anti-*Cosmos,* a television series called *Space, Time, and God* that was to counter Sagan's "conspicuously materialist approach to the universe." The program never got off the ground, but its premise survives: that there are two ways to think about science. You can be a theist, believing that behind the veil of randomness lurks an active, loving, manipulative God, or you can be a materialist, for whom everything is matter and energy interacting within space and time. Whichever metaphysical club you belong to, the science comes out the same.

In the hands of as fine a writer as Gingerich, the idea almost sounds convincing. "One can *believe* that some of the evolutionary pathways are so intricate

HELIX NEBULA, also known as the eye of God.

NASA/ESA/STScI/SCIENCE PHOTO LIBRARY

and so complex as to be hopelessly improbable by the rules of random chance," he writes, "but if you do not believe in divine action, then you will simply have to say that random chance was extremely lucky, because the outcome is there to see. Either way, the scientist with theistic metaphysics will approach laboratory problems in much the same way as his atheistic colleague across the hall."

Thus, a devoutly Christian geneticist such as Francis S. Collins, author of *The Language of God* and leader of the Human Genome Project, can comfortably accept that "a common ancestor for humans and mice is virtually inescapable" or that it may have been a mutation in the *FOXP2* gene that led to the flowering of human language. The genetic code is, after all, "God's instruction book."

But what sounds like a harmless metaphor can restrict the intellectual bravado that is essential to science. "In my

view," Collins goes on to say, "DNA sequence alone, even if accompanied by a vast trove of data on biological function, will never explain certain special human attributes, such as the knowledge of the Moral Law and the universal search for God." Evolutionary explanations have been proffered for both these phenomena. Whether they are right or wrong is not a matter of belief but a question to be approached scientifically. The idea of an apartheid of two separate but equal metaphysics may work as a psychological coping mechanism, a way for a believer to get through a day at the lab. But theism and materialism don't stand on equal footings. The assumption of materialism is fundamental to science.

Richard Dawkins, in *The God Delusion,* tells of his exasperation with colleagues who try to play both sides of the street: looking to science for justification of their religious convictions while evading the most difficult implications—the existence of a prime mover sophisticated enough to create and run the universe, "to say nothing of mind reading millions of humans simultaneously." Such an entity, he argues, would have to be extremely complex, raising the question of how it came into existence, how it communicates—through spiritons!—and where it resides.

Dawkins is frequently dismissed as a bully, but he is only putting theological doctrines to the same kind of scrutiny that any scientific theory must withstand. No one who has witnessed the merciless dissection of a new paper in physics would describe the atmosphere as overly polite.

Sagan, writing from beyond the grave (actually his new book, *The Varieties of Scientific Experience,* is an edited version of his 1985 Gifford Lectures), asks why, if God created the universe, he left the evidence so scant. He might have embedded Maxwell's equations in Egyptian hieroglyphs. The Ten Commandments might have been engraved on the moon. "Or why not a hundred-kilometer crucifix in Earth orbit?... Why should God be so clear in the Bible and so obscure in the world?"

He laments what he calls a "retreat from Copernicus," a loss of nerve, an emotional regression to the idea that humanity must occupy center stage. Both Gingerich and Collins, along with most every reconciler of science and religion, invoke the anthropic principle: that the values of certain physical constants such as the charge of the electron appear to be "fine-tuned" to produce a universe hospitable to the rise of conscious, worshipful life.

But the universe is not all that hospitable—try leaving Earth without a space suit. Life took billions of years to take root on this planet, and it is an open question whether it made it anywhere else. To us carboniferous creatures, the dials may seem miraculously tweaked, but different physical laws might have led to universes harboring equally awe-filled forms of energy, cooking up anthropic arguments of their own. SA

George Johnson is author of Fire in the Mind: Science, Faith, and the Search for Order *and six other books. He resides on the Web at talaya.net*

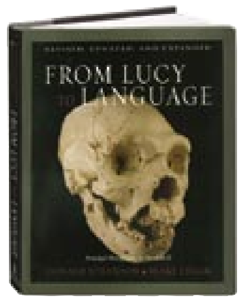
Editors' note: Two other noteworthy books on religion by scientists have appeared recently: E. O. Wilson's The Creation: A Meeting of Science and Religion *(W. W. Norton, 2006) and Joan Roughgarden's* Evolution and Christian Faith: Reflections of an Evolutionary Biologist *(Island Press, 2006).*

THE EDITORS **RECOMMEND**

FROM LUCY TO LANGUAGE

by Donald Johanson and Blake Edgar. Revised, updated and expanded. Simon & Schuster, 2006 ($65)

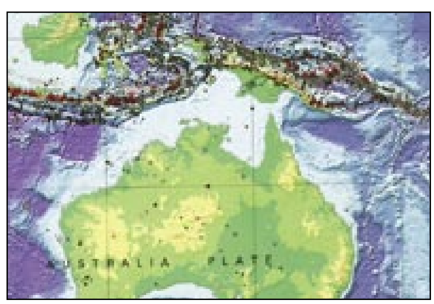
Since the original edition was published in 1996, paleoanthropologists have made several important finds. Among them are *Sahelanthropus tchadensis,* a seven-million-year-old specimen uncovered in Chad that has features that are part ape, part hominid, and *Homo floresiensis,* diminutive people who apparently were not *Homo sapiens* and who lived on the Indonesian island of Flores as recently as 13,000 years ago. Johanson, director of the Institute of Human Origins and best known for his discovery of the "Lucy" skeleton, and Edgar, a writer and an editor at the University of California Press, present other new finds and add updates throughout the book. With more spectacular photographs by David Brill, most of them depicting specimens at actual size, the new tome is even more awe-inspiring than the earlier version.



THIS DYNAMIC PLANET: WORLD MAP OF VOLCANOES, EARTHQUAKES, IMPACT CRATERS, AND PLATE TECTONICS

Smithsonian Institution, USGS and U.S. Naval Research Laboratories, 2006 ($14)

Another new edition of a classic is this map, which shows the dynamic plate tectonic processes that shape the planet. All elements of the updated map are digital, and an interactive version is at www.minerals.si.edu/minsci/tdpmap. The site is somewhat slow, but the enormous amount of data at your fingertips is worth the patience required. You can make your own regional map, for example, by choosing the layers of information you want (volcanoes, plate motion, latitude and longitude, and so on). The one-by-1.5-meter paper version is a bargain at $14.



Drawing to an Inside Flush

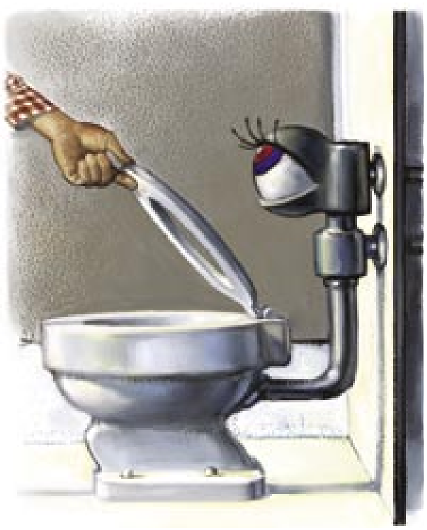
A TALE OF TWO TOILETS **BY STEVE MIRSKY**

In early August a short item crossed my desk about troubles on a movie set in Mumbai, formerly known as Bombay. Actors and crew were trying to film a scene in a public restroom for the Bollywood blockbuster *Keep at It, Munnabhai*. But when the actors walked past the autoflush urinals, they inadvertently set off the sensors. The water would noisily flow, and the scene would go down the drain. "At one point, with so many unit members inside the loo, all the flush sensors went berserk and started flushing simultaneously," recounted Raju Hirani, the film's director, according to the Associated Press. "We actually had to vacate the loo briefly to stop the urinals from flushing."

The flushing toilets of Mumbai (officially ranked as the 14,287th Wonder of the World, by the way, just behind the Hanging Gardens of Piscataway but before the Colossus of Killiecrankie by the A9 road) took me back to my own misadventure with automatically flushing toilets. This escapade took place at personal computing's headwaters, the headquarters of Microsoft.

The year was 1997: Researchers publicly announced the existence of Dolly the cloned sheep, *The Simpsons* passed *The Flintstones* as the longest-running animated television series ever, and newfangled autoflush toilets were helping America stay hygienic. It was a heady time.

I was at the annual meeting of the American Association for the Advancement of Science, held that February in Seattle. Journalists attending the conference were also invited to visit Microsoft near Redmond. So I went and inevitably had to go. I left a lecture (which I recall was about the ongoing efforts to create reliable voice-recognition software, just to give you an idea of the scope of that still unsolved problem) and wandered until I found a bathroom. I entered, put the seat up and proceeded as usual. After which, being committed to the commonweal of my fellow fellows, I tried to flush. And was thwarted at every turn.



While searching in vain for a handle or button or even dangling chain, I noticed a small, dark rectangle in the middle of which was a luminous red dot. I knew then that I was in the presence of electronic technology.

Clearly, this object was a sensor designed to automatically flush the toilet once the end user zipped away. And yet no flush would gush, no surge would purge, no swirl unfurled. I refused to leave the room before disposing of all the evidence, so I began a meticulous debugging analysis. And through a careful consideration of the geometries, relative positions and functions of all the objects in the setup, I concluded that the sensor had been located in a place where it could be blocked by only one thing—the upraised seat. With the seat up, the system was convinced that a request was still being processed. So I put the seat down.

That simple act, the savior of millions of marriages, solved the problem. What we used to call "the electric eye" suddenly was alerted to the fact that the task was complete. Water began its flow to the sea, and another wee aliquot of processed caffeine started its journey to Puget Sound.

As I washed up, I reflected that the situation at Microsoft was probably explicable in one of two ways. One possibility was that whoever installed the componentry had used state-of-the-art motion-sensor technology along with deep ergonomic theory and application to trick men into putting down the toilet seat. The other option was that they had accidentally cobbled together a Rube Goldbergian arrangement that in effect replaced the old-fashioned toilet-seat handle with the seat itself. As a Windows user who has to click "Start" to turn my computer off, and as a man who knows that most men wouldn't try all that hard to flush in the first place, I'm betting on the latter. SA

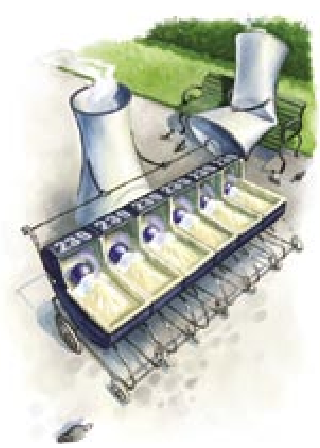
ILLUSTRATION BY MATT COLLINS; PHOTOGRAPH BY FRANK VERONSKY

How do fast breeder reactors differ from regular nuclear power plants?

P. Andrew Karam, a senior health physicist with MJW Corporation in Amherst, N.Y., explains:

In addition to generating power, as regular nuclear plants do, fast breeder reactors utilize so-called fast neutrons to produce up to 30 percent more fuel than they consume.

Nuclear reactors generate energy through fission, the process by which an atomic nucleus splits into two or more smaller nuclei. During fission, a small amount of mass converts into energy, which a reactor harnesses by a controlled chain reaction; when a uranium nucleus splits, it produces two or more neutrons that other nuclei absorb, causing them to undergo fission as well. As more neutrons are released in turn, continuous fission occurs. Fission produces neutrons with high energies that move extremely quickly. These fast neutrons do not efficiently cause fission. To maintain the reaction, most reactors use a moderator, commonly water or helium, to slow the neutrons to optimum energies.



In contrast, a fast reactor uses a less effective coolant, such as liquid sodium, so the neutrons remain high energy. Although these fast neutrons are not as good at causing fission, they provide a side benefit: they are readily captured by the isotope uranium 238, which then becomes plutonium 239. Certain reactor designs maximize plutonium production for use as fuel or to produce nuclear weapons; reactors that can make more fuel than they consume are called breeders. They do this by exploiting the natural preponderance of uranium 238, which does not fission readily, over uranium 235, which does.

Commercial nuclear reactors typically use uranium fuel that has been enriched so it is up to 8 percent uranium 235. Although this minority isotope does the bulk of the fissioning to produce energy, most of the atoms in the fuel are uranium 238—potential future plutonium 239 atoms.

Plutonium 239 happens to be even better at fissioning than uranium 235, but this fissioning reduces the plutonium content of the fuel. This is why many breeder reactors are fast reactors. Fast neutrons are ideal for plutonium production because they are easily absorbed by uranium 238 to create plutonium 239 yet cause less fission than slower-moving neutrons.

Creating extra fuel in nuclear reactors is not without its concerns: the plutonium produced could contribute to nuclear proliferation, and the fuel reprocessing necessary for plutonium extraction creates radioactive waste and potentially high radiation exposure. For these reasons, President Jimmy Carter halted such reprocessing in the U.S. in 1977. Today only India, Russia, Japan and China have operational fast breeder programs; the U.S., the U.K., France and Germany have effectively shut down theirs.

What do butterflies do when it rains?

Michael Raupp, professor of entomology at the University of Maryland, offers this answer:

What butterflies ultimately do in the rain is avoid it. A storm is no trivial matter—for a 500-milligram monarch butterfly, getting hit by a 70-milligram raindrop would be equivalent to you or me being pelted by water balloons with twice the mass of small bowling balls.

Storms may also hinder a butterfly's mobility. In preparation for flight, these aerial acrobats expose their wings to direct sunlight, which rapidly warms their flight muscles. Overcast skies block the solar radiation they need to take wing.

Thus, when the sky darkens, butterflies seek shelter in their roosts: protected locations such as tall grasses or leafy plants. But when sunshine returns, they often resume patrolling and courting within minutes. So the next time thunder rumbles, take a cue from the butterflies: find shelter, but as soon as the sun comes back, go out and enjoy. SA

For a complete text of these and other answers from scientists in diverse fields, visit www.sciam.com/askexpert

MATT COLLINS